Optimizing the measurement scales of qualitative variables to improve bankruptcy prediction

A thesis submitted in partial fulfilment of the requirements

for the Degree

Of Masters of Commerce in Accounting

At the University of Canterbury

By Piumini Bowatte

2019

University of Canterbury

Department of Accounting and Information Systems

Senior Supervisor: Associate Professor Dr. Michael Falta

Co-Supervisor: Dr. Julia Wu

Abstract

I extend Altman's (1968) multivariate linear discriminant bankruptcy model by adding two variables, auditor-client familiarity (AFT) and senior management turnover (SMT). These two variables have not been considered in the bankruptcy modelling, albeit the literature suggests them to relate to firm performance. The addition of the two qualitative poses the question as to i) which measurement scale (nominal, ordinal or interval) should be used, and ii) how to determine the optimal scale distances such that the variables can be included into the regressions with the intention to improve their bankruptcy forecasting ability.

I collect financial and non-financial information for 70 UK construction companies to test above research questions. My final 7-variables model predicts 69-out-of-70 firms of my sample correctly into healthy and failed firms (as a comparison, the re-estimated Altman 5-financial ratios model predicts 56-out-of-70 firms correctly).

My thesis was motivated by the recent Carillion PLC collapse in the UK that attracted wider attention from the media and government. The case was portrayed as a surprise event, yet several models in my study indicate a strong decline in performance over 3 years prior to their collapse.

Acknowledgements

I would first like to thank my senior supervisor Associate Professor Dr. Michael Falta of the Department of Accounting and Information Systems at the University of Canterbury. I am gratefully indebted to his invaluable time and continuous guidance throughout my research. I would also like to thank my Co-Supervisor Julia Wu, as well as all the lecturers at the University of Canterbury who provided their professional knowledge to add to this research.

Secondly, I would like to thank my friends, work colleagues and my church family for their endless support and encouragement throughout my studies.

Finally, I must express my profound gratitude to family and my Godparents for providing me with unfailing support and continuous encouragement throughout my study at UC.

I extend my gratitude to everyone, this accomplishment would not have been possible without you.

Piumini Bowatte, August 2019

Table of Contents

Chapter 1 Introduction

Corporate bankruptcy prediction is a major research domain in accounting and finance. Prediction models are used to assess a company's creditworthiness and its ability to continue operating. Users of the prediction models are investors, professional brokers, financial institutions, banks and other stakeholders. Knowing something about the level of financial distress a firm is in will help stakeholders making better investment and lending decisions. Financial distress models can also signal if and to what degree a corporation has recovered from a downturn.

Prediction models do not capture all the modes of company failure – it depends on what variables a model contains, and what data have been used to train (estimate) the parameters of the model. There are different reasons for company failure, and financial ratios alone may or may not predict the creditworthiness of a firm. In practice, bankers and financial analysts use financial and non-financial information to measure company performance (e.g. Bloomberg news feeds). When a large corporation collapses, blame for the firm's failure is often directed towards the senior management and the auditors: for example, the collapse of Carillion PLC, one of the UK construction giants, has drawn wide media attention. Three of the big four audit firms were involved with Carillion. Because of the seeming failure of the audit function and the economic impact of the Carillion bankruptcy, it was not just a discussion that (re)emerged, but actual political debate over the break-up of the big four audit firms. Some of the headlines read "Carillion: accountants accused of 'feasting' on company (Davies, 2018). BBC Business news report the UK Competition and Markets Authority CMA "… recommends accountancy market overhaul" and "… stopped short of calling the Big Four accountancy firms to be broken up." ("Comeptition watchdog recommends accountancy market overhaul," 2019). Marriage (2018) also report the suggestion by Financial Reporting Council that urged to break-up the big four in the UK and make them audit-only firms (Marriage, 2018). The collapse also questioned the senior management and there were claims made by MPs who stated that the "board was too busy stuffing their mouths with gold to care about workers and should be banned from running other firms" (Wearmouth, 2018).

My research was motivated by the collapse of Carillion PLC. From a research perspective, I wondered if and how far in advance the collapse could have been flagged. My main result, depicted in Figure 1-1, is that early warning signals about Carillion's (low) performance were available as early as 2006, and in 2015, as an investor, I would have sold my shares.

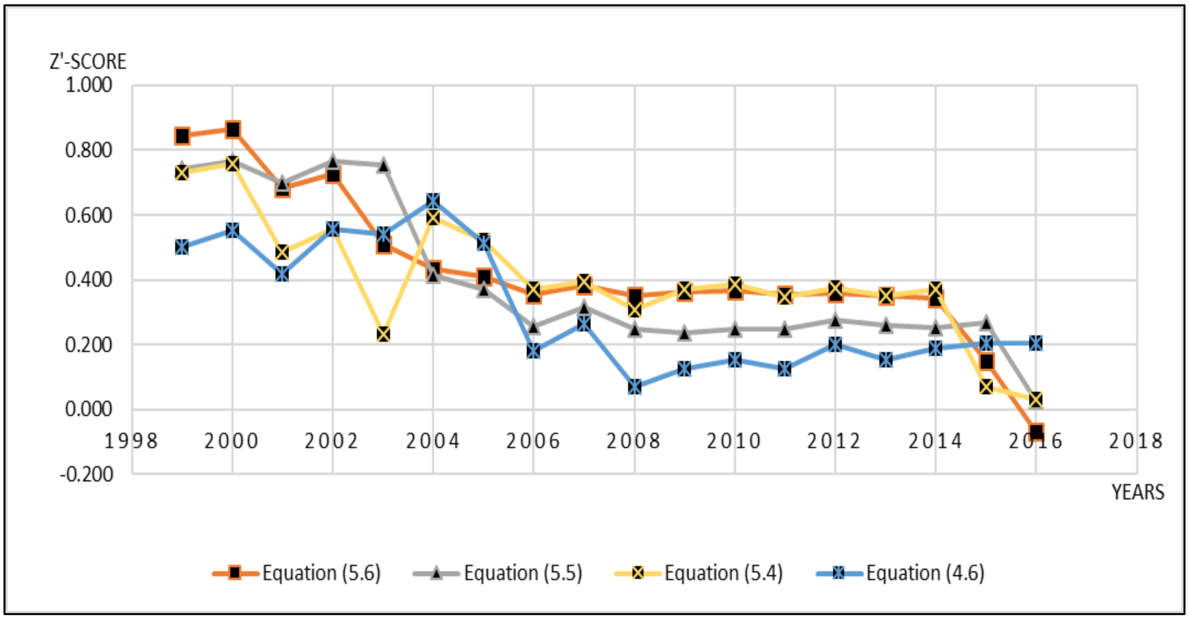


Figure 1-1 **Carillion's performance over time**. Measurements using four regression models.

The model with the best predictive ability shown in Figure 1-1 (orange data) includes variables that have not been considered in the bankruptcy literature as yet: the Carillion collapse suggests to investigate the merits of including variables to bankruptcy models that contain information about auditors and senior management. The research question therefore is: do qualitative variables improve the predictive accuracy of bankruptcy models?

The two qualitative variables I have used are Auditor-client Familiarity Threat (AFT) and Senior Management Turnover (SMT). I have added the two variables to Altman's (1968) Z-score, which is a weighted index of financial ratios, and tested for its predictive accuracy using a sample of 70 UK construction entities. I chose the Z-score due to its popularity in the literature (Bellovary, Giacomino, &

Akers, 2007). My research then focuses on measurement issues of the non-financial, ordinary scale variables, AFT and SMT. I use Monte Carlo simulations to determine an optimal measuring scale for AFT and SMT.

My research adds to the bankruptcy literature. I suggest adding qualitative variables to discriminant analysis models because they improve overall prediction accuracy.

The rest of the paper is structured as follows. Chapter 2 reviews previous literature on the Z-score and the use of qualitative variables therein. Chapter 3 describes the research methodology, Chapter 4 my initial data analysis, and in Chapter 5 I analyse the properties of the qualitative variables by running Monte Carlo simulations. Chapter 6 uses the derived results from earlier chapters to assessing the performance of Carillion PLC. Chapter 7 looks critically into the presentation of the Z-score model in accounting text books. Finally, Chapter 8 summarises the project and gives an overview of further research.

Chapter 2 Literature Review

The bankruptcy prediction literature is highly methodological. Most studies use financial ratios to predict corporate failure, creditworthiness and ability to continue in operation**.**

Seminal bankruptcy studies that made use of ratio analysis were done in the 1930's. Later, Beaver (1966) used an univariate analysis (single factor analysis) to predict corporate failure in the US. He analysed 14 financial ratios that potentially signal company failure. His univariate analysis sets the stage for multi-variate models, such as Altman (1968) Z-score which is a multivariate linear discriminant (MDA) model that uses five financial ratios. Altman used data from the 1950ies in a matched sample of 66 failed and non-failed US manufacturing firms. Deakin (1972) showed in an alternative prediction model using MDA by replicating the Beaver (1966) study that using the originally suggested 14 financial ratios together would improve the model.

The bankruptcy literature is further characterised in that studies uses different variables and different samples. Thus, it is unclear which of the models produced in the literature would apply best in what situation. Generally, the common feature of the models is that they used financial ratios which measure profitability, liquidity, efficiency, solvency, leverage and operational activity.

1.1 The use of MDA

Altman's Z-score is the benchmark model to beat if a researcher proposes their own model. The Z-score thus remains widely popular in the literature (Bellovary et al., 2007). Based on the Z-score, Altman developed further versions, such as the Z' for private sector manufacturing firms, Z'' and ZETA® for service sector firms (Altman, 2000).

After Altman Z-score was published, there were more MDA models re-estimated for non-US data. These MDA models used a combination of Altman Z-score variables as well as new financial ratios. Taffler (1982) used the Z-score and re-estimated financial ratios for UK manufacturing companies using six variables and an unmatched sample of firms (23 failed and 45 non-failed). Ferner and Hamilton (1987) re-estimated the Z-score for New Zealand manufacturing firms. Edmister (1972) tests the usefulness of financial ratio analysis for small businesses using the Altman Z-score and the Beaver (1966) study. Mason and Harris (1979) developed the Z-score using eight completely different financial ratios and applied it on a sample of UK construction firms. Laitinen (1992) used MDA for predicting the failure of newly founded firms.

The use of new variables, samples and the country specific settings developed into a main theme in the bankruptcy literature. Bellovary et al. (2007) analysed 165 bankruptcy prediction models, including the MDA approaches. Albeit the focus of the research continues to be in developing new prediction models the authors suggest that future research ought to improve existing model rather than putting effort into new model developments. In relation to MDA prediction models, one of the critiques is that they use predominantly financial variables. The data for these variables come from annual reports. Thus there is a limitation on the frequency of observations, and the financial ratios alone may not capture some firms' failure if the reasons for poor performance lies outside of the information content of these variables. This is where non-financial information may enhance the predictive ability of company distress.

1.2 The use of qualitative variables

There is only a limited amount of literature that uses non-financial variables for predicting corporate failure. According to du Jardin (2009) who analysed 190 published articles between the 1960s and 2008, only thirteen per cent considered non-financial information. Those studies that did consider non-financial variables included firm-specific characteristics such as its operational environment, leadership characteristics and age of the firm.

Mostly and for large firms, failure is not a sudden event that happens within a day, say, but it is a result of the effects of many events and decisions (Argenti, 1976). One of the models that uses non-financial ratios is the A-score produced by Argenti (1976). The concept behind the A-score is that distress occurs due to management defects and mistakes that were made. Abidali and Harris (1995) show that distress can be measured better when both the Z- and A-scores side by side to predict financial failure. However, adding non-financial variables to the MDA models has not been widely adopted since.

The importance of non-financial information such as managerial decision making has been recognised in academia elsewhere. Recently, Pervan and Kuvek (2013) test the predictive power of non-financial variables by comparing two logit models. One is using financial ratios and the other model ius using a combination of financial and non-financial variables. The second model resulted in a higher accuracy rate (88.1%) than the first model (82.8%). They used an unmatched sample of 825 client firms in which 698 are non-failed and 127 are failed. The non-financial variables included, for example, firm age, firm size and auditor opinion. Assaad (2010) added two non-financial variables (audit quality and downside risk variables) to the logit model which resulted in a better prediction performance with consistent and robust testing using estimation and test samples. His final model consisted of twelve variables and predicted well up to five years prior to bankruptcy.

In summary, the MDA and logit bankruptcy literatures suggest that adding non-financial information to prediction models improves their performance. However, little is known about AFT and SMT in relation to bankruptcy studies, I therefore review the characteristics and use of these two non-financial variables elsewhere.

1.3 The SMT variable

Senior management in a firm has the decision-making power on options and therefore is responsibility for a firm's performance. The senior management refers to senior executives, the chairman, directors and presidents. Senior management's responsibility is to act in the best interest of the owners of the firm, i.e., investors and lenders. If the top managers involve themselves in fraudulent activities and show lack of integrity to external parties of the firm, severe consequences such as bankruptcy may arise. One of the most often used examples to this effect is the downfall of Enron in the USA.

(Gilson, 1989) defines SMT as any change in top management in a given year. Gilson then investigated SMT in distressed firms and found that in a sample of 381 US companies, 52% of them experienced high SMT. Earlier, Schwartz and Menon (1985) similarly found that 45% out of 126 US bankrupted firms had a high level of executive turnover. Furtado and Karan (1990) state that managers are likely to be removed from a firm that is close to bankruptcy or soon after entering liquidation proceedings. The top management turnover indicates uncertainty that then impacts on a firm's performance. Kaplan (1993) also suggests that poor performance of a firm increases with increasing SMT. Agrawal, Jaffe, and Karpoff (1999) indicate that management turnover whether voluntarily or involuntarily induced is a major event for the firm and will influence the future performance of the management. Darrat, Gray, Park, and Wu (2016) sampled 217 bankrupt firms and 9100 healthy firms in the US for the period of 1996 to 2006 and researched the relationship between corporate governance characteristics and financial distress. They find that large boards and a higher percentage of inside directors decrease the possibility of company failure. Further they suggest that the firm is better off with a longer standing CEOs. These findings were consistent with Hsu and Wu (2014) who did a UK-based study and find that firms with greater proportions of independent directors are less likely to fail.

The above studies suggest that there is a positive correlation between firm performance and SMT because high SMT signals company uncertainty, and lower CEO turnover consistency perhaps. The above literature also suggests that examining board composition and the change in board composition may be a worthwhile exercise for future variable identification and selection in bankruptcy research.

1.4 The AFT variable

There were different approaches to measuring the association between corporate failure and audit independence. Past researches commonly used auditor opinion (Blay, Geiger, & North, 2011; Chen & Church, 1996; Sikka, 1992) and the non-audit services by the audit firms (Firth, 1981; Hudaib & Cooke, 2005; Hussey, 1999).

External auditors are supposed to provide reliability and credibility to the financial reports of their clients through an independent audit procedure. Therefore, auditors' independence is important and comes to question when there are big audit failures such as Enron and WorldCom (US), or Carillion (UK). Auditor independence is defined as the ability to act with integrity and objectivity (Hudaib & Cooke, 2005). International Standard on Auditing 200 (ISA 200, UK) states that auditor independence safeguards the audit firm to issue an audit opinion without being affected by influences that may compromise their opinion. Independence is therefore meant to enhance an auditor's ability to act objectively and to maintain a professional skepticism (ISA 200.A.18).

One of the threats to the objectivity and independence of the audit is familiarity threat, which is the degree of closeness between the auditor and her client. Closeness is understood in terms of duration and/or coziness between the two agents. The familiarity threat and independence has been researched and measured using qualitative measures such as matched interviews, surveying the financial directors audit partners and investors (Dart, 2011) and there are measures done in previous literature observing the familiarity threat and audit tenure (Carey & Simnett, 2006; Hussey, 1999). Early literature by Firth (1981) found that longer audit-client relationships reduce audit fees as well as increased audit expertise. In contrast, Knapp (1991) text-based survey suggests that longer audit tenures decrease the audit quality. In other words, length of audit engagement affects auditor independence and increases the familiarity threat. (Dart, 2011; Firth, 1981; Hussey, 1999) found that there was no association between the length of the audit tenure and the independence in appearance to the public, especially not to the investors. Hussey (1999) surveyed 265 private UK firms of which 119 firms had a long-term association with the audit firms (more than ten years): although 44.9% display a high level of AFT, investors were more concerned about non-audit services given to the clients.

There are arguments that favour audit-tenure. One of which is that an auditor-client long-term relationship may make the audit procedure more coherent and less time consuming because the audit firm has the knowledge and expertise about the firm (see Firth, 1981; Hussey 1999). Then again, a number of studies including DeFond and Park (2001), Carey and Simnett (2006), Arrunada and Paz-Ares (1997), Hoyle (1978) suggest that long-term relationships with the client have a tendency of predictable about the audit opinion results rather than having the auditor be alert, critical and independent with their analysis. Also, (Carey & Simnett, 2006; McLaren, 1958) who investigate the long-term relationships between auditors and their clients show that loyalty develops. Loyalty however impacts on the auditor's independence and objectivity which works against fulfilling of ISA 200, for example. Recent research by Ghosh and Tang (2015) observed that auditor resignations have a positive association with firm distress levels. Adams, Krishnan, and Krishnan (2017) found a negative correlation between the size of a client and auditor resignation.

The above discussion suggests a link between audit opinion and distress modelling. However in my research and for my sample of companies, audit opinion did not lead to an improvement of prediction performance. Instead, I then have focused on AFT that appears to have an at least indirect effect on an auditor's sharpness to report objectively. This will apply to both well-running and poorly performing firms, however, the effects from the lack of objectivity will be more severe for the latter. Adams's et al. (2017) result means that if the client firm is large there is a smaller chance for auditor resignation, in which cases AFT increases. While Adams et al. (2017) is not a bankruptcy study, Gosh and Tang's (2015) study relates audit resignation with firm failure, and because any audit resignation will terminate the audit-client relationship, it will also determine the size of AFT. Lindahl (1992) identified a direct relationship between financial distress and auditor change. Beattie and Fearnley (1995) report on a variety of other reasons for auditor change and refer to a wide literature on consequences including negative market reactions.

In summary, the link between auditor resignation, AFT and financial distress merits an investigation. Particularly, into the usefulness of AFT to carry information content that improves forecast ability of financial distress.

Chapter 3 Methodology

3.1 Econometric model

The statistical model chosen in this research is MDA. The general form of MDAs is as follows,

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n + \mathcal{E}, \quad (3.1)$$

where Z' is the discriminant score, β_0 is the intercept, β_i (*i=1,2..,n*) are the discriminant coefficients, X_i (*i=1,2,…,n*) are the independent predictor variables, and Ɛ is the error term of the discriminant function.

Altman (1968) reports on the first MDA bankruptcy prediction model. MDA is a statistical technique used to classify observations into one of several groups depending upon distinct characteristics of choice. In bankruptcy studies this choice is either failed (1) or non-failed (0). The dependent variable is thus a binary variable. MDA is based on several assumptions. It assumes dichotomous data, Normality of the error term, equal dispersion of the variance-covariance matrix across the two groups, and the absence of multicollinearity (Balcaen & Ooghe, 2006).

Altman's (1968) Z-score combines five ratios into a single weighted index. In order to estimate the MDA that represent the Z-score, a matched sample of 33 failed and 33 non-failed manufacturing firms in the years 1946 to 1965 were chosen. Initially, 22 ratios were considered representing liquidity, leverage, solvency, profitability and performance ratios. From these, all but five were eliminated to yield the following formula:

$$Z_{Altman} = 1.0\,STA + 1.4\,RETA + 3.3\,EBITTA + 0.6\,MVETL + 1.2\,WCTA \quad (3.2)$$

where, STA is Sales over Total Assets, RETA is Retained Earnings over Total Assets, EBITTA is EBIT (Earnings Before Interest and Tax) over Total Assets, MVETL is Market Value of Equity over Total Liability, and WCTA is Working Capital over Total Assets. If Z > 2.99 a firm is classified healthy and if Z

< 1.81 a firm is classified as failed. If the Z-score is in between 1.81 and 2.99 then a firm cannot be classified to either group.

Altman later revised his model to include private companies by replacing the MVETL variable with BVETL (Book value of equity/Total liabilities). Thus the Z'-model is:

$$Z'_{Altman} = 0.998\ STA + 0.847\ RETA + 3.107\ EBITTA + 0.420\ BVETL + 0.717\ WCTA. \qquad (3.3)$$

Here, if Z' > 2.90 then the company is classified as healthy and if Z' < 1.23 the company is classified as failed. If Z' is between 1.23 and 2.90 then no group assignment is possible.

For the purpose of my research I will be using Z'_{Altman} as my baseline model because my sample consist of both private and public companies. Also note that Z'_{Altman} does not contain an intercept β_0, an issue discussed below. Furthermore, Altman chose two cut-off values. For example, the number 1.81 in relation to Z_{Altman} is the lowest Z-value of a healthy company, however, not the highest value of all of the failed companies (which is 2.99). Therefore, the more Z-values lie in between these two numbers, the less useful the model becomes. I therefore use only one cut-off value.

3.1.1 Cut-off value determination

Cut-off values classify the discriminant score into failed and non-failed groups. According to Hair (1998) the optimal cut-off score differs depending on whether the sample is matched or unmatched. I use a matched sample, and I determine only one cut-off point according to Equation 3.4. Thus,

$$\text{Cut-off} = \frac{1}{2}(\overline{Z}'_{failed} + \overline{Z}'_{non-failed}) \qquad (3.4)$$

where,

$$\overline{Z'}_{failed} = \frac{1}{n}\sum_{i=1}^{n} Z'_{i,fail} \tag{3.5}$$

$$\overline{Z'}_{non-failed} = \frac{1}{n}\sum_{i=1}^{n} Z'_{i,non-fail.} \tag{3.6}$$

3.1.2 Determination of model performance

The norm in bankruptcy studies is to judge model performance using Type I and Type II errors. A Type I error is a false positive, and a Type II error is a false negative. Thus, these errors depend on how the Null hypothesis is formulated. If, for example, H_0 is companies are predicted as non-failed, then a Type I error is that the company is predicted as failed while the company is non-failed. A Type II is when the company is predicted as a non-failed while in fact the company fails. Using this approach is insensitive to the particular reasons for failure and therefore explore an alternative approach to measuring model performance below, starting in Section 4.2.

3.2 Data

The financial data I obtained through the Orbis database of Bureau Van Dijk (A Moody's analytical Company). Orbis captures information on around 300 million companies' across the globe; however, financial information was available for approximately 30 million companies. More than 99 per cent of companies covered in this database are private companies. The Orbis database provides a standardised version of the financial statements. This database provides a company's history, activities, business lines, news, and original filing (local registry filing). It also provides information on the company advisors, auditors, and senior managers, both previous and current.

Another database used in gathering data is Endole. Endole is a UK online database that provides comprehensive and authentic company data from government bodies. It has 7 million company profiles. However, this online database is not subscribed by the University of Canterbury. But Endole offers free subscription which I made use of. The free Endole subscription provides an overview of the company key people and company documents such as the group accounts, auditors filing, date of incorporation, and insolvency filings. These documents are extracted from the UK Companies House[1]. I have used the Endole database mainly to gather the non-financial data on auditors, AFT, and senior management, SMT.

3.3 Sample

My sample consist of large public and private construction sector companies in the UK. I gathered data for the non-failed and failed companies for the period between 1991 and 2017. A large firm means herein to have more than 250 employees or more than £50 million in annual turnover. Construction sector firms were selected using the Statistical classification of Economic Activities in the European Community, commonly referred to as NACE (for the French term **n**omenclature statistique des **a**ctivités économiques dans la **C**ommunauté **e**uropéenne) codes. NACE codes 41 to 43 (primary codes) belongs to the construction sector.

I have chosen a matched sample because the previous literature that uses MDA predominantly does so (Altman, 1968, 2000; Altman, Iwanicz-Drozdowska, Laitinen, & Suvas, 2017).

The initial screening resulted in 344 non-failed firms and 67 failed firms. Based on data availability (Orbis database excludes the companies with no recent financial data and public authorities) I obtain a sample of 35 failed firms all of which provide financial information for all three years prior to failure. I then randomly matched these with 35 non-failed firms. The matching criterion is total assets: I first

[1] Companies House is the UK official website that register company information and make it available to the public. Companies house incorporate and dissolve limited companies and has all the online filing information. https://www.gov.uk/government/organisations/companies-house

calculated the average total asset for the failed companies (£760 million) using the 1-year prior failure financial information. Two and three years prior to failure, these companies had larger total assets. Therefore, I choose 35 healthy firms which are in the vicinity and tangentially higher of the £760 million, i.e. in the range between £1 billion and £760 million.

The final sample of the 70 matched observations consists of 71% private and 29% public companies, as shown in Table 3-1.

Table 3-1 **Private and Public firm distribution in the sample (*N=70)***

	sample (n=70)	failed	Non-failed
Private	50	31	19
Public	20	4	16

I acknowledge the slight dominance of private over public firms which was expected. Regarding the sample composition, I refer to Altman (1983) who provides for a precedence study in using a mix of private and public firms to estimate Z'_{Altman}.

3.3.1 Definition for failed firms

Failed companies are defined in legal terms using the UK insolvency law. The UK Insolvency Act 1986 regulates companies which are unable to repay their debt. While UK bankruptcy law concerns the rules for natural persons, the term insolvency is used for companies formed under the Companies Act 2006. If a company fails to meet its debt obligations, the company could go into either administration or liquidation. The administration procedure can result in either survival or liquidation as a result of company asset conversion to cash or cash equivalents. Liquidation is an insolvency procedure where the liquidator converts (or sells) the company assets and distributes the funds to the creditors. There are two modes of liquidation: Compulsory liquidation and Voluntary Liquidation. Compulsory liquidation is the liquidation of a company by court order, and voluntary liquidation is a non-court based procedure to liquidate a company where creditors or board members are involved in the liquidation procedure. My study includes firms that entered into the liquidation process either through administration, voluntary or compulsory liquidation.

The Orbis database has five categories for potentially active firms (active, the default of payment, receivership, dormant, and branch) and seven categories for inactive firms that no longer carry out business activities (bankruptcy, dissolved, dissolved-merger, dissolved-demerger, in liquidation, branch, and no precision). Among these categories, I only choose the ‘active’ category to select non-failed firms. For the failed firms the UK definition for liquidation (see above). The applicable Orbis categories are therefore the ‘liquidation’ and ‘dissolved’ categories.

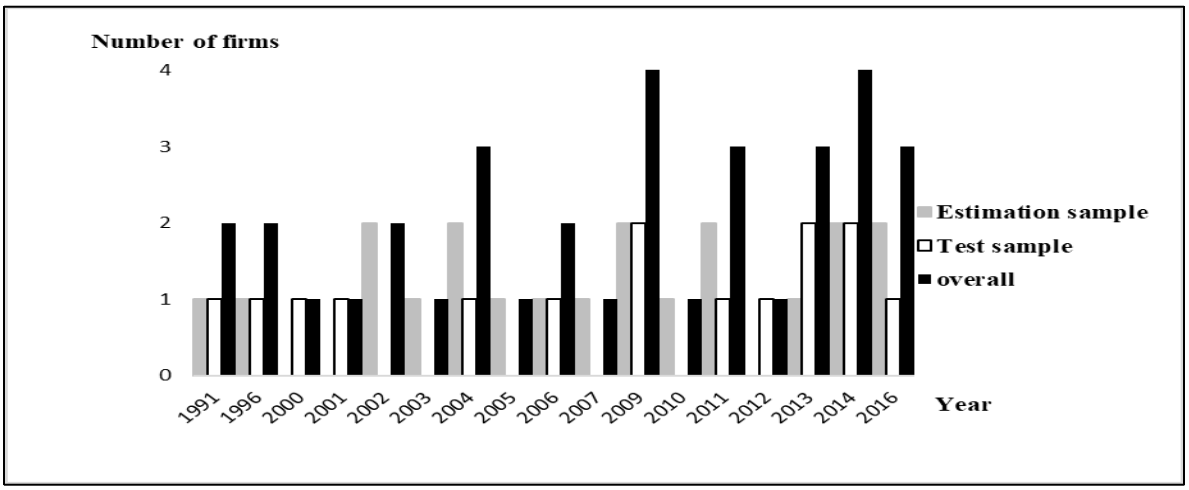


Figure 3-1 **Sample of failed companies (N=35)**

Figure 3-1 displays the distribution of my full sample into estimation and test samples, and for failed firms only. Note that years without any observation are not displayed. The horizontal scale therefore is not equally spaced. From 1991 to 1999 there are only 4 companies that failed thus, my sample mainly consists of failed companies between the years 2000 to 2016. High failure rates fall between 2006 to 2009 and 2011 to 2014. The high failure rates in these two periods is suspected to be due to the Global Financial Crisis of 2008 and a recession period in the UK construction sector, respectively. According to the UK construction sector index, construction sector firm failure was high due to the recession in 2010 and 2011: “The recession has hit construction extremely hard and the forecast recovery is likely to be long and slow.”

Mark Farrar, Chief Executive of Construction Skills (as quoted by Hinchliffe, 2010) . Therefore, I am not surprised to see the high failure rates between the two periods in my sample. My sample represents the UK construction sector economic trends.

3.4 Qualitative Variables

The addition of qualitative variables are at the heart of my research. These two qualitative variables are added to test whether they improve bankruptcy prediction accuracy. As stated in Chapter 2, there is only a limited amount of literature on the use of non-financial information in failure prediction models. The two qualitative variables that I introduce are AFT and SMT. Both variables end up to be ordinal, whereas the Z'_{Altman} five variables are ratio scale variables. Issues from mixing ratio and ordinal variables inside an MDA model have not been widely researched to date, albeit there are of course no distributional conditions posed on any independent variables in an OLS regression. The bigger problem is about determining the distances for the ordinal variables which are unknown, per se, and to complicate matters, will also have to be selected in relation to the ratio scale variables. In Chapter 5, the research setup to find an appropriate measuring scale is explained and implemented.

3.4.1 Measuring AFT

On 6th of October 2009 the UK Auditing Practices Board issued a revised Ethical Standard 3 (ES3) which specifies audit engagement (ICAEW, 2019). The revised ES3 states that the maximum audit engagement partner rotation is five years and with a minimum of five years not involved in the audit afterwards. In the UK, this standard applies for an audit commencing on or after the 15th December 2009 (ICAEW, 2019). My sample data are from 1991 to 2017, and contains the pre revision of ES3 and post revision of ES3.

The International Ethics Standards Board for Accountants defines AFT as follows in their 2012 report. “A familiarity threat is the threat that due to a long or close relationship with a client will be too sympathetic to their interest or too accepting of their work” (IESBA, 2012, p. 3).

Because of regime and definitional changes throughout my sample period, and the distinction between audit partner and audit firm, I first group AFT-related observations as follows:

- G1: post-ES3 which had the maximum partner rotation reduced to 5 years;
- G2: pre-ES3 which is characterised by a 7 years maximum partner rotation rule;
- G3: audit firm resignation.

Note that in the UK, an audit firm may, and usually would, remain with their clients, particularly large clients. Audit partner rotations apply to changes in engagement roles from within an audit firm.

In a second step, I introduce order between groups G1 to G3. Based on the literature, AFT is lower for G1 than it is for G2 because of the shorter allowable engagement period. As for G3, although an auditor resignation abruptly ends the familiarity threat (it resets AFT to zero for the new auditor), however precedence in my choice of order is that the very nature of the event has been shown to directly relate to financial distress. Also, audit firm resignation is unusual in the UK because of the rotation system, which suggests that if it happens, the conflict is beyond a personal level. Thus G3 ranks most negatively on the ordinal AFT scale.

Table 3-2 **AFT groups.**

	Group		
	AFT G1	AFT G2	AFT G3
Scenario	After ES3 revision (Audit partner rotation is maximum of 5 years)	Before ES3 revision (Audit partner rotation is longer than 5 years)	Audit firm resignation
overall	43	8	19
Failed	9	8	18
Non-failed	34	n/a	1

Table 3-2 outlines the three groups (and scenarios) and ranks them. The table also shows the distribution of my sample into the three categories. It is not surprising to see most of the non-failed firms in AFT G1. On the other hand, the number of failed firms is approximately uniformly distributed across all three groups. Although AFT G3 does weakly relate to the concept of AFT, it signals strongest company financial distress.

In a third step, I have assumed value ranges for the ordinal AFT scale. The numbers and ranges chosen, reflect a qualitatively-based decision which was informed by the average size of the financial ratios and proportionality contributions of the terms in a Z-score model. The value ranges are displayed in Table 3-3.

Table 3-3 **Valuation range for each AFT group.**

	AFT G1			AFT G2			AFT G3		
Range	0.9	to	0.3	0.3	to	-0.3	-0.3	to	-0.9

Generally, an observation in G1 will add to the Z-score, which increases the companies' classification chance as a healthy firm. On the other hand, an observation falling into G3 will reduce the Z-score and thus increase the chance for that company to be classified as a failed company. Chapter 5 is dedicated to the determination of an optimal choice for the AFT scale.

3.4.2 Measuring SMT

In the UK, when a senior manager is terminated, the company must file a TM01 application. If several director are terminated, separate TM01s have to be filed. The filing of the TM01 documents that contain director terminations appear at least 6 months after the event. This does not affect missing observations today in 2019 when by sample period ended in 2017. I have obtained all TM01 for my sample companies through the Endole database. The TM01 contain the dates of director termination which I then can associate with a particular financial year (and data).

These applications for the limited companies are available digitally through the UK Companies House website and Endole online database. Three scenarios were observed in UK Companies House online database which I put into the following three groups:

- G1: no director is being terminated (low risk).
- G2: one director terminated or replaced (indicates risk);
- G3: two or more directors are terminated (leadership uncertainty).

In comparison to the AFT variable, the SMT scale order is straight-forward. G1 associates with the lowest probability of financial distress, which increases for G2 and G3, respectively. Table 3-4 gives an overview.

Table 3-4 **SMT groups.**

	Group		
	SMT G1	SMT G2	SMT G3
Scenario	No director terminated	one director terminated	two or more director terminations
overall	43	8	19
Failed	9	8	18
Non-failed	34	n/a	1

Table 3-4 shows that there is a noticeable difference in the director turnover between failed and non-failed companies. Average director termination in the failed groups is 4. In the case for non-failed companies director termination is evenly split between none and one. Gilson (1989) found that there is a high SMT in failed companies which corresponds to my data in Table 3-4.

The number of director terminations is a number. But in forming three SMT groups, I now have created a ordinal scale for which I now need inter-group distances. I follow the same procedure and line of argumentation as for the AFT variable which yields the valuation ranges as shown in Table 3-5.

Table 3-5 **Valuation range for each SMT group.**

	SMT G1	SMT G2	SMT G3
Range	0.9 to 0.3	0.3 to -0.3	-0.3 to -0.9

SMT G1 is given a strictly positive range as the no directors terminated means that there is a long standing leadership. SMT G2 is given both negative and positive range as is some risk involved in replacing one director. SMT G3 is given a strictly negative range of values as it indicates high uncertainty in the leadership when two or more directors are terminated.

Chapter 4 Data analysis Part I: Estimation of Z'_{Altman} model

This chapter contains the initial data analysis of the Z'_{Altman} model. I have split, randomly, my sample of 70 firms into a matched estimation sample (*N=40*, i.e. 20 failed and 20 non-failed firms) and matched test sample (*N=30*). I then re-estimate the Z'_{Altman} model, with and without an intercept, with my estimation sample. For all failed firms, I use data 1 year before bankruptcy, and I match these by year with data from the non-failed firms. The re-estimated models are then analysed using statistical significance tests, misclassification rates and checks for error term Normality.

4.1 Financial ratio analysis

The variables used in the original Altman models (3.2) and (3.3) are shown in Table 4-1. Each variable captures a different aspect of a company that reflects on its performance.

Table 4-1 **The original five ratios used in the Z'_{Altman} model.**

Variable	*Definition of the ratio*	*Type of ratio*
STA	Sale / Total Assets	Efficiency ratio
EBITTA	Earnings Before Interest and Tax / Total Assets	Profitability ratio
RETA	Retained Earnings / Total Assets	Leverage and efficiency
BVETL	Book Value of Equity / Total Liabilities	Leverage ratio
WCTA	Working Capital / Total Assets	Liquidity ratio

STA is an efficiency ratio that measures the firm's ability to generate sales from the assets under their control. The EBITTA ratio measures how effectively the firm uses its assets to generate earnings. The RETA ratio measures the cumulative profitability over time as a proportion of total assets. The BVETL ratio measures how much the company's assets can decline in value before the liabilities exceed the assets and the firm becomes insolvent. WCTA is the liquidity ratio that measures the firm's ability to pay short-term liabilities.

These five financial ratios are used to re-estimate model (3.3). Firstly, I provide in Table 4-2 descriptive statistics for the five ratios using my estimation sample.

Table 4-2 Descriptive statistics for the five financial ratios in the failed and non-failed groups of my estimation sample. Panel A shows the non-failed and failed group statistics and Panel B shows the change in failed and non-failed group statistics. Mean is the sum of the observations divided by the total number of observations. Standard Error indicates how close the sample mean from the population mean. Median is another measure of central tendency, which is calculated by ordering the data from the lowest to the highest. The median is the number in the middle. Standard Deviation is the square root of the variance. Sample variance measures the dispersion of the data from the mean (SV = sum of (X-mean of X)^2/ (number of observation - 1). Kurtosis focuses on the tails of the distribution. Skewness measures the asymmetry of the data. The Range is the difference between the largest and smallest value (Max-Min). Δ is the change between failed and non-failed descriptive statistics.

	STA		RETA		EBITTA		BVETL		WCTA	
Descriptive statistics	Non-failed	Failed	Non-failed	Failed	Non-failed	Failed	Non-failed	Failed	Non-failed	Failed
Mean	1.96	1.44	0.22	-0.10	0.08	-0.10	0.55	0.06	0.27	-0.02
Change in the Mean	0.52		0.32		0.18		0.49		0.29	
Median	1.86	1.37	0.17	0.09	0.07	-0.03	0.32	0.13	0.18	-0.01
Minimum	0.65	0.00	0.05	-1.79	0.01	-0.65	0.08	-0.63	-0.01	-1.56
Maximum	4.72	4.40	0.61	0.33	0.16	0.16	2.00	0.69	0.73	0.46
Standard Error	0.25	0.26	0.04	0.11	0.01	0.05	0.13	0.07	0.05	0.09
Standard Deviation	1.10	1.18	0.16	0.50	0.04	0.21	0.60	0.31	0.24	0.40
Sample Variance	1.20	1.39	0.03	0.25	0.00	0.04	0.36	0.10	0.06	0.16
Kurtosis	0.98	0.74	0.96	6.82	-0.32	1.13	1.70	0.51	-0.85	12.22
Skewness	1.02	1.02	1.34	-2.51	0.38	-1.34	1.65	-0.43	0.72	-3.07
Range	4.07	4.40	0.56	2.11	0.14	0.81	1.92	1.32	0.75	2.02

Table 4-2 shows that size of and difference between the means for the failed and non-failed groups varies considerably. Especially RETA, EBITTA and WCTA seem to have high between-group discriminatory potential because the non-failed group has a positive mean and the failed group has a negative mean. However, WCTA has a high kurtosis for the failed firm with 12.22, which means that the data set has heavy tails will make this variable to have a more volatile impact on the Z-scores.

To illustrate the potential to discriminate between failed and non-failed companies, I show the ratio distribution for EBITTA in Figure 4-1. Considering that none of the ratios for the non-failed companies is negative, and all but 5 are negative for the failed companies, a univariate bankruptcy analysis would obtain an expected prediction accuracy of approximately 100% - (5/40)% = 87.5%. Based on the estimation sample properties, this variable thus suggests to have discriminatory power between failed and non-failed firms.

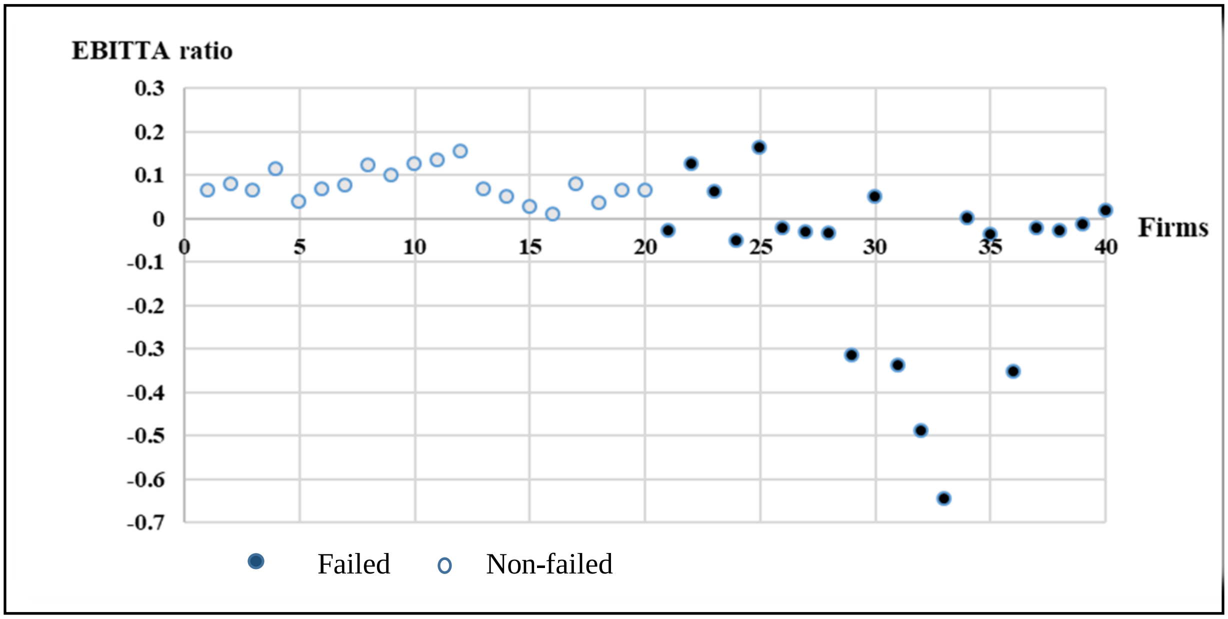


Figure 4-1 **EBITTA ratio distribution for the estimation sample (N=40).**

One of the assumptions of the MDA is absence of multicollinearity between the independent variables. Multicollinearity is when predictor variables are strongly correlated. I used tolerance and the Variance Inflation Factor (VIF) in order to test the collinearity between the variables. The tolerance and VIF are measured as follows,

$$Tolerance = 1 - R_j^2 \qquad (4.1)$$

$$VIF = \frac{1}{Tolerance} \qquad (4.2)$$

where R_j^2 is the coefficient of determination of the regression of predictor j on all the other predictors. A generally accepted rule is if the tolerance value is less than 0.1, and the VIF value is above 10, there is multicollinearity between the variables (Alin, 2010).

Table 4-3 **Collinearity statistics**. VIF and the tolerance.

	Dependent Variables (dv)											
	Z		STA		RETA		EBITTA		BVETL		WCTA	
Independent Variables	VIF	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF	Tolerance
STA	1.16	0.87	**dv**	**dv**	1.37	0.73	1.16	0.87	1.08	0.93	1.13	0.89
RETA	5.87	0.17	5.55	0.18	**dv**	**dv**	3.82	0.26	5.74	0.17	2.82	0.34
EBITTA	2.10	0.48	2.10	0.48	2.44	0.41	**dv**	**dv**	2.05	0.49	1.84	0.55
BVETL	2.49	0.40	2.33	0.43	1.09	0.91	0.41	0.41	**dv**	**dv**	2.22	0.45
WCTA	4.49	0.22	4.38	0.44	2.16	0.46	3.92	0.26	3.99	0.25	**dv**	**dv**

Table 4-3 shows the collinearity statistics using various predictor variables as dependent variables. No one of the regressions displays a VIF that is higher than 10 and a tolerance factor below 0.1. I thus conclude that there is no significant multicollinearity between the variables and I can use these five financial variables to (re)estimate the Z'_{Altman} model.

4.2 Estimation model

The Z'_{Altman} model (3.3) is now re-estimated. Using my estimation sample and MS Excel 2016 for the regression analysis, I obtain:

$$Z'_{estimate} = 0.17STA - 1.05\,RETA + 2.1\,EBITTA + 0.36BVETL + 0.87\,WCTA \qquad (4.3)$$

which compares to a reprinted original model (3.3) in which I drop the Altman suffix for the remainder of my thesis,

$$Z' = 1.0\,STA + 0.85\,RETA + 3.11\,EBITTA + 0.42\,BVETL + 0.72\,WCTA \qquad (4.4)$$

as follows: The coefficient estimates for each independent variable have changed. Noticeably the sign of RETA has changed. Clearly, the differences arise from the use of two different samples.

One change that should be noted is that the original Z' does not have an intercept, which Altman states is due the choice of the particular statistical software (Altman, 2000). The question therefore arises, what effect the exclusion of the intercept has on the coefficient estimates, and the predictive ability of a model. These issues will be discussed in the next section.

4.2.1 Intercept in the Z' models

The more general MDA regression has a form as follows,

$$Z'_{+\alpha} = \alpha + \beta_1\, STA + \beta_2\, RETA + \beta_3\, EBITTA + \beta_4\, BVETL + \beta_5\, WCTA + \varepsilon \qquad (4.5)$$

where α is the intercept, Ɛ is the error term and the β_i (*i =1, 2, 3, 4, 5)* are the parameter of interest. The intercept is the expected mean value of the Z' score when all of the independent variables are 0. The corresponding estimated model using my estimation sample yields

$$Z' = Z'_{+\alpha} = -0.01 + 0.18\, STA - 1.6\, RETA + 2.1\, EBITTA + 0.4\, BVETL + 0.9\, WCTA. \qquad (4.6)$$

Comparing these parameter estimates with their equivalents in Equations (4.3) and (4.4) further displays the variation in sign and magnitude, now not only due to the sample of choice, but also in dependence of the intercept.

In order to illustrate the issue of the intercept in a Z-score model, I have taken the EBITTA variable and my estimation sample (*N=40)* to demonstrate the effect of an intercept. The regression equations are as follows,

$$Z''_{+\alpha} = \alpha + \beta_1 EBITTA + \varepsilon \qquad (4.7)$$

$$Z''_{(\alpha=0)} = \beta_2 EBITTA + \varepsilon \qquad (4.8)$$

where $Z'_{+\alpha}$ and $Z'_{(\alpha=0)}$ are the discriminant variables that take on the values of 0 and 1, $\beta_{1,2}$ are the correlation coefficients, α is the intercept and Ɛ is the error term. The regression results are displayed in Figure 4-2.

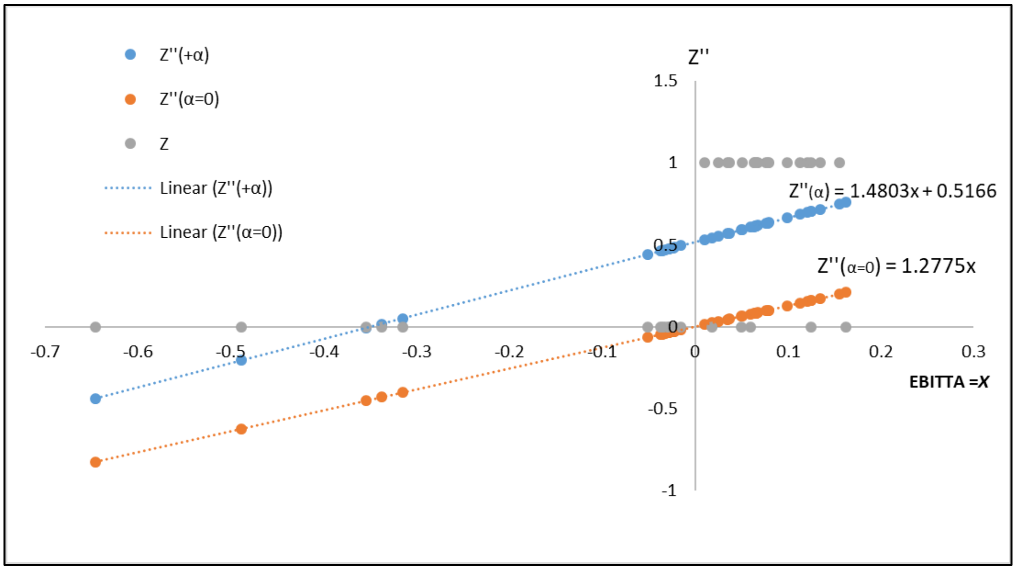


Figure 2-2 **The effect of an intercept.**

In Figure 4-2 the two regression slopes are different, as expected. From visual inspection alone, the imposed restriction on the $Z''_{(\alpha = 0)}$ model to go through the origin obtains a worse fit with the grey Z-scores. Whether or not adding another degree of freedom (intercept) will allow for better prediction performance is now an empirical research question addressed below.

4.2.2 Predictive performance of the Z' model with and without the intercept

I now analyse the predictive performance of Equations (4.3) and (4.6) using both estimation (*N=40*) and hold-out samples (*N*=30). The predictive performance of the two models is measured by the number of misclassifications (cf. Chapter 3.1.2). The results are displayed in Table 4-4.

Table 4-4 **Type I and Type II error analysis:** for both test and estimation samples for Equations (4.3) and (4.6). Accuracy is calculated by ((N-total error)/N)

	Test sample (N=30)		Estimation sample (N=40)		Total	Accuracy for all
	Type I	Type II	Type I	Type II	Error	firms N=70
Equation 4.6	5	7	4	5	14	70%
Equation 4.3	5	8	4	5	15	69%

Overall, model $Z'_{+\alpha}$ makes one less Type II error in the test sample and thus has a slightly higher overall predictive accuracy of 70% in comparison to the Z'_{est} model (4.3).

The question arising is: yes, depending on the sample, but otherwise, can we conclude that the intercept-including model (4.6) is superior? The commonly applied predictive comparison is based on Type I and Type II errors; however, can we truly conclude that model (4.6) is better than model (4.3)? For example, it is unclear if the 5 Type I errors made by models (4.3) and (4.6) in the test sample misclassify the same five firms. If a different sample was chosen, perhaps the results would tell otherwise. Ono (2018) compared two models by Altman (1968) and by Ohlson (1980) and suggested that observing the classification pattern on an individual firm basis would allow a more robust conclusion about model performance. It was found, for example, that the number of observations used in estimation and test samples will change which of the two models predicts better.

Using the firm-level classification table shown in Figure 4-3 now confirms that in my case, both models commit the classification errors with the same firms, apart from the one more error in model (4.3).

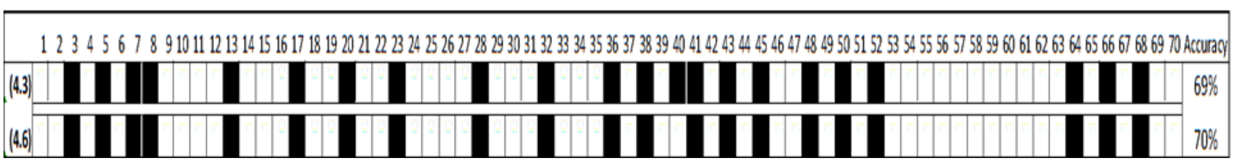


Figure 4-3 The individual misclassification of all 70 firms in the sample.

Thus the Z' model (4.6) will be my baseline model from here on. Next I perform a further analysis around the regression condition of error term Normality.

4.3 Error term Normality

An error term in a regression equation represents the variations in the dependent variable that the independent variables do not explain. In order for the coefficient estimates on the independent variables to be unbiased, the average value of the error term must equal to zero and approximately be Normally distributed. If the error terms are not Normally distributed suggests that systematic information to explain the variation in the dependent variable is included in it, i.e., we speak of an omitted variables problem.

In this Section the error term Normality is tested on the Z' model using the Shapiro-Wilk test and Normal Q-Q plots using the IBM SPSS Statistics 25 software package.

4.3.1 Shapiro-Wilk test

The Shapiro-Wilk test tests the Null hypothesis that a sample came from a Normally distributed population. The test statistics is as follows,

$$W = \frac{(\sum_{i=1}^{n} a_i \, x_{(i)})^2}{\sum_{i=1}^{n}(x_i - \mu)^2} \qquad (4.9)$$

where n is the number of observations, x_i is the sample of size n, $x_{(i)}$ is the sample sorted in increasing order, and a_i are the coefficients. If the p-value that represents the significance of W is less than 0.05, the Null hypothesis is rejected, which means that there is significant evidence that the data are not Normally distributed.

The Shapiro Wilk test results for the Z' model yields a p-value of 0.290 which is above the p-value of 0.05; thus it cannot reject the null hypothesis that the error term is Normally distributed. The Z'_{est} model thus passes the assumption of the error term Normality.

4.3.2 Normal Q-Q Plots

A Normal Q-Q plot suggests that if the distribution of the data is Normal, then the points on the Normal Q-Q plot would lie approximately on the Y = X line. Normal Q-Q plots compare two Normal probability distributions by plotting their quantiles against each other. For the error term, it compares the observed with the expected error terms. For illustration purposes, the following example shown in Figure 4-4 is provided, where a Normally distributed sample was used. The random sample used consists of 1000 randomly generated Normally distributed numbers in a Normal Q-Q plot as well as in a histogram.

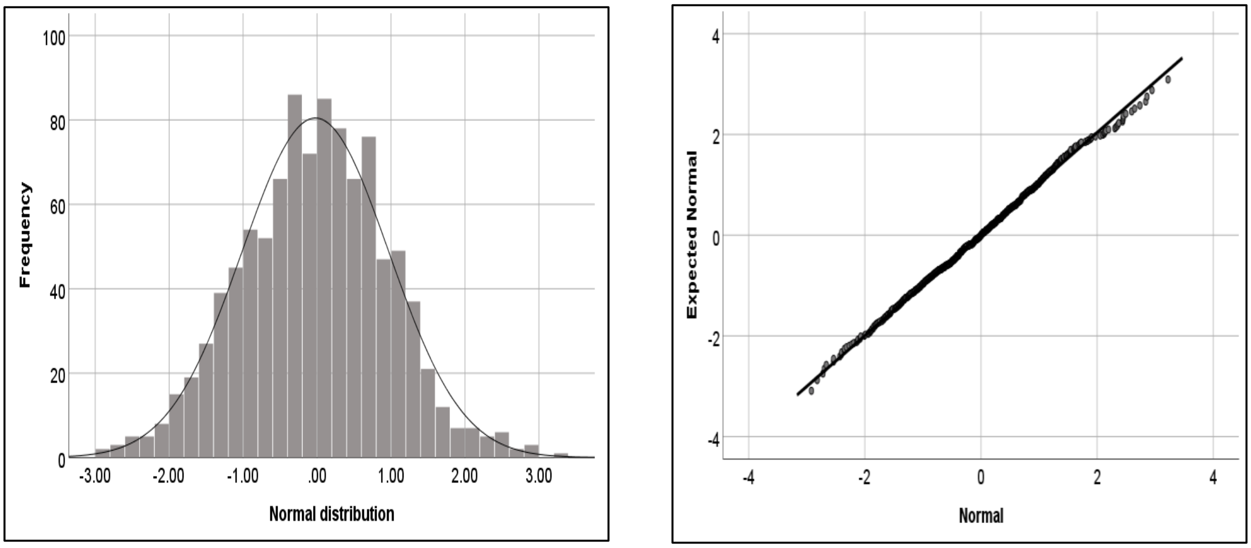


Figure 4-4 Illustration of Normal distribution and associate Normal Q-Q plot.

Figure 4-5 compares the Normal Q-Q plot of the Z' model error normality curve.

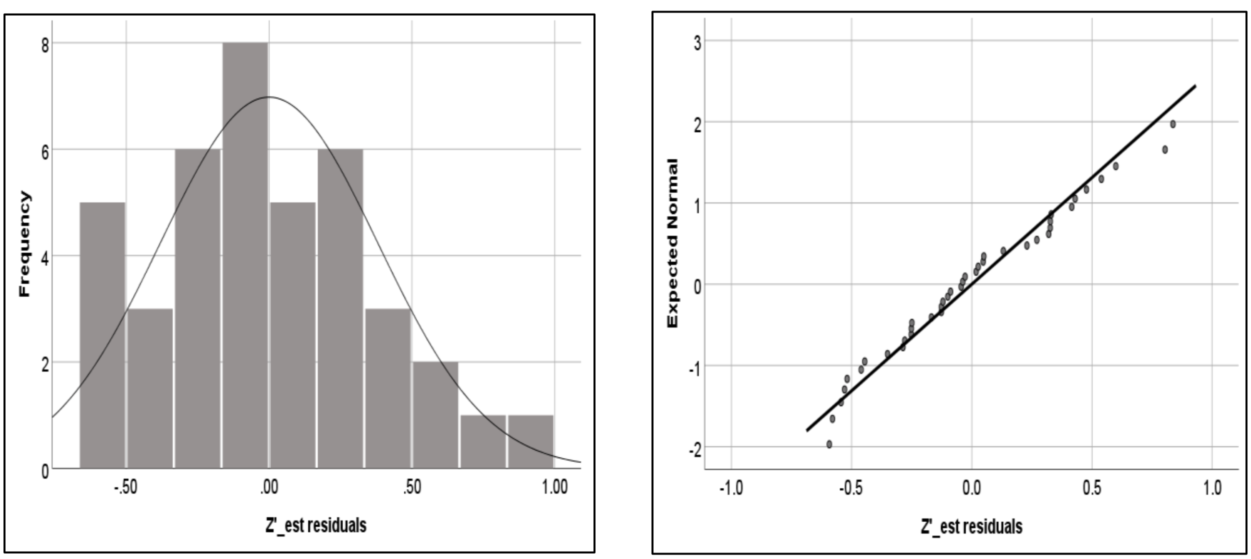


Figure 4-5 Error term distribution for the Z′ model using histogram and Normal Q-Q plot.

The Normal Q-Q plot shows that the points are close to the Y=X line, and the histogram resembles a symmetric, 0-centered Normally-shaped distribution.

From this visual inspection of the information presented in Figure 4-5 and in comparison to the optimal case shown in Figure 4-4. In summary, both the Shapiro-Wilk test (Section 4.3.1) and the Normal Q-Q plot (this section) indicate that the Z' model error term is approximately Normally distributed.

4.3.3 The robustness of error Normality in the Z' model

Further analysis re error Normality is run on the Z' model by excluding one of the five independent variables. Because the missing information is now contained in the error term, I expect that the reduced models do not (easily) pass the error term Normality tests. I thus analyse the following regressions:

$$Z' = \alpha + \beta_1 \text{(STA)} + \beta_2\text{(RETA)} + \beta_3\text{(EBITTA)} + \beta_4\text{(BVETL)} + \beta_5\text{(WCTA)} \qquad (4.10)$$

$$Z'_{(-STA)} = \alpha + \beta_2 \text{(RETA)} + \beta_3\text{(EBITTA)} + \beta_4\text{(BVETL)} + \beta_5\text{(WCTA)} \qquad (4.10a)$$

$$Z'_{(-RETA)} = \alpha + \beta_1 \text{(STA)} + \beta_2\text{(EBITTA)} + \beta_3\text{(BVETL)} + \beta_5\text{(WCTA)} \qquad (4.10b)$$

$$Z'_{(-EBITTA)} = \alpha + \beta_1 \text{(STA)} + \beta_2\text{(RETA)} + \beta_4\text{(BVETL)} + \beta_5\text{(WCTA)} \qquad (4.10c)$$

$$Z'_{(-BVETL)} = \alpha + \beta_1 \text{(STA)} + \beta_3\text{(EBITTA)} + \beta_4\text{(BVETL)} + \beta_5\text{(WCTA)} \qquad (4.10d)$$

$$Z'_{(-WCTA)} = \alpha + \beta_1 \text{(STA)} + \beta_2\text{(RETA)} + \beta_3\text{(EBITTA)} + \beta_4\text{(BVETL)} \qquad (4.10e)$$

The corresponding coefficient estimates are displayed in Table 4-5.

Table 4-5 **Coefficient estimates for Equations 4.10 to 4.10x, x=a,..,c.**

		Coefficients				
		STA	RETA	EBITTA	BVETL	WCTA
Equation	α	β1	β2	β3	β4	β5
4.10	-0.01	0.18	-1.06	2.10	0.36	0.87
4.10a	0.35	n/a	-0.68	2.05	0.22	0.59
4.10b	0.14	0.12	n/a	1.33	0.24	0.20
4.10c	0.01	0.17	-0.73	n/a	0.33	1.02
4.10d	0.16	0.13	-0.83	1.96	n/a	1.05
4.10e	0.08	0.13	-0.43	2.43	0.45	n/a

The corresponding error term Normality test statistics using the Shapiro-Wilk approach are shown in Figure 4-6.

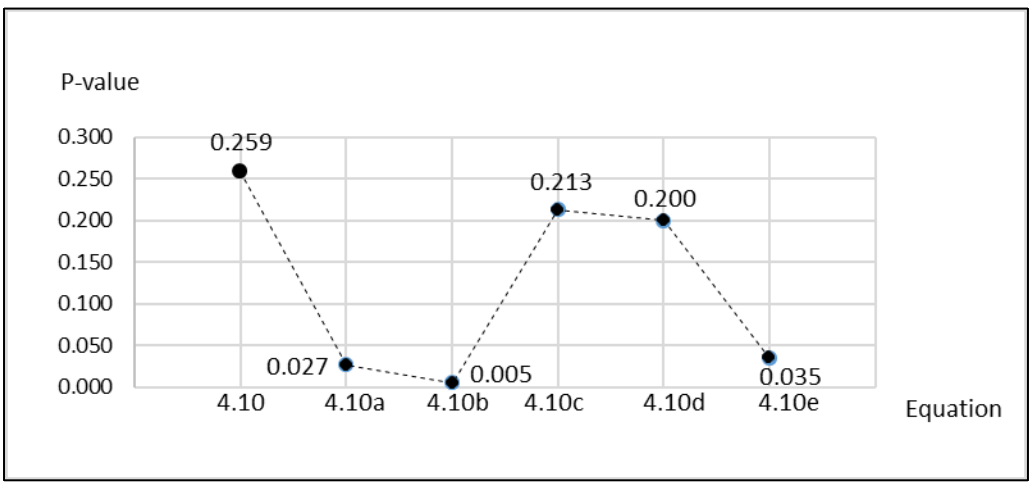


Figure 4-6 **Error term Normality statistics for Equations 4.10x, x=a,..,e.**

Figure 4-6 clearly shows deterioration of the error term Normality test statistics for the reduced models in comparison with the 5-variables Z' model. In particular, the error Normality assumption is rejected in Equations 4.10a, 4.10b and 4.10e; in Equations 4.10c and 4.10d the error term Normality is worsened in relation to Equation (4.10). I therefore conclude it not to be useful to exclude any of the five original ratio variables.

4.4 Summary

The five ratios capture different weaknesses in the companies. The Z' model passes some of the most important statistical regression assumptions. The inclusion of an intercept is important. However, it still make a high number of misclassification errors. This suggests that there is an omitted variables problem, not so much based on statistical assumptions, but in the ability to capture firm failure modes which have not been discriminated for by the five financial ratios.

This, in combination with the rational developed in the literature review, then justifies the reason to include two further (qualitative) variables, AFT and SMT.

Chapter 5 Data analysis Part II: Qualitative variable analysis

The focus in this chapter is on issues around including AFT and SMT into the Z' model. This firstly includes issues about the design of measurement scales for the AFT and SMT variables. AFT and SMT are ordinal variables which implies that the distances between the groups are undetermined. Because I add them to a bankruptcy index (Z-Score) that contains rational scale variables (financial ratios), the chosen in-between-groups distances will i) in/decrease the Z-score accordingly, and ii) influence the parameter estimates on the other financial ratios when the model is estimated. For example, if the in-between-group distances are chosen to be very small, little to no discriminatory contribution from the AFT and SMT variables will be achieved. On the other hand, if the in-between-group distances are chosen to be very large, they will nullify the contributions of the other (financial) information with respect to discriminating poorly from well performing firms.

Thus, I use the estimated Z' model (4.6) and compare its predictive performances with the following three models:

$$Z'_{SMT} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_7 SMT \quad (5.1)$$

$$Z'_{AFT} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_6 AFT \quad (5.2)$$

$$Z'_{AFT,SMT} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_6 AFT + \beta_7 SMT \quad (5.3)$$

The predictive performance of models (5.1), (5.2) and (5.3) will depend on the choice of measurement scale, as argued above. The research question that I am faced now therefore is: *What are the optimal in-between-group distances when ordinal scale variables are added to a discriminant function that contains ratio-scale independent variables*?

To my knowledge, the issue arising from a mixed variable discriminant analysis has not been discussed in the bankruptcy literature to date. In fact, the discipline of statistics has only recently made some advances with respect to such research questions (e.g., Wissmann and Toutenburg, 2007). The determination of a particular scale may be based on theoretical or pragmatic considerations. I have chosen the pragmatic approach and implemented a Monte Carlo simulation analysis to determine useful (optimal) in-between-group distances.

5.2 Design of measurement scales

Following Section 3.4, both ordinal variables have been grouped according to their impact as shown in Table 5-1, Row 2.

Table 5-1 **Qualitative grouping of AFT and SMT.**

Groups	G1	G2	G3
Impact	Positive effect for the company; low risk in company failure	Both positive and negative impact in company; medium risk in company failure.	Negative impact in company; high risk of company failure
Range	+0.9 to +0.3	+0.3 to -0.3	-0.3 to -0.9

Guided by the size of the estimated parameter coefficients in model (4.6), I have chosen the value ranges for each group, as shown in Row 3 of Table 5-1. The values allow for symmetric and asymmetric weighting options, and the range margins allow for minimum and maximum distances of 0.6 and 1.8, respectively. I have chosen to investigate 136 different combinations that represent a wide variety of measurement alternatives (cf. Figure 5-1 and Table A-1 given in the Appendix).

As hypothesised in the introduction of this chapter, I expect to see a correlation between the model success rates and the distances chosen for the scales of the SMT and AFT variables. Thus on top of the 136 alternatives, I test a few extreme alternatives in order to i) gauge the degree with which the prediction sensitivity among the 136 alternatives alters, and ii) if I have chosen large (small) enough range values (+0.9 and -0.9). To test point ii), I additionally chose seven points between -5 to +5, and for small scale values, I chose five points between -0.01 and 0.4. Note that a zero distance indicates that the model is run without the ordinal variables, i.e., models (5.1), (5.2) and (5.3) revert to model (4.6).

Overall, I thus run 148 AFT and SMT simulations, using MS Excel, that each estimate the regression parameters using my estimation sample. Based on the estimated models (5.1), (5.2) and (5.3) I then can determine the corresponding prediction accuracies which are shown in Figure 5-2.

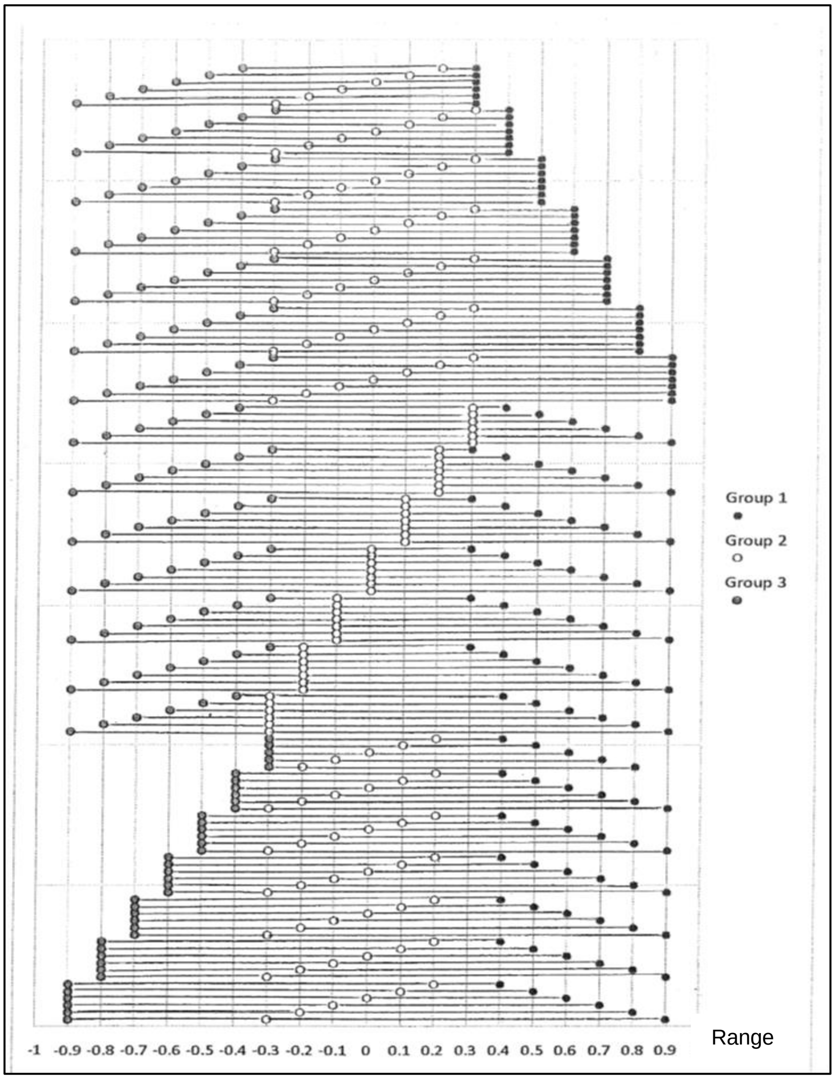


Figure 5-1 **Scaling options for Monte Carlo simulation.** 136 cases the ordinal variables AFT and SMT will be measured on.

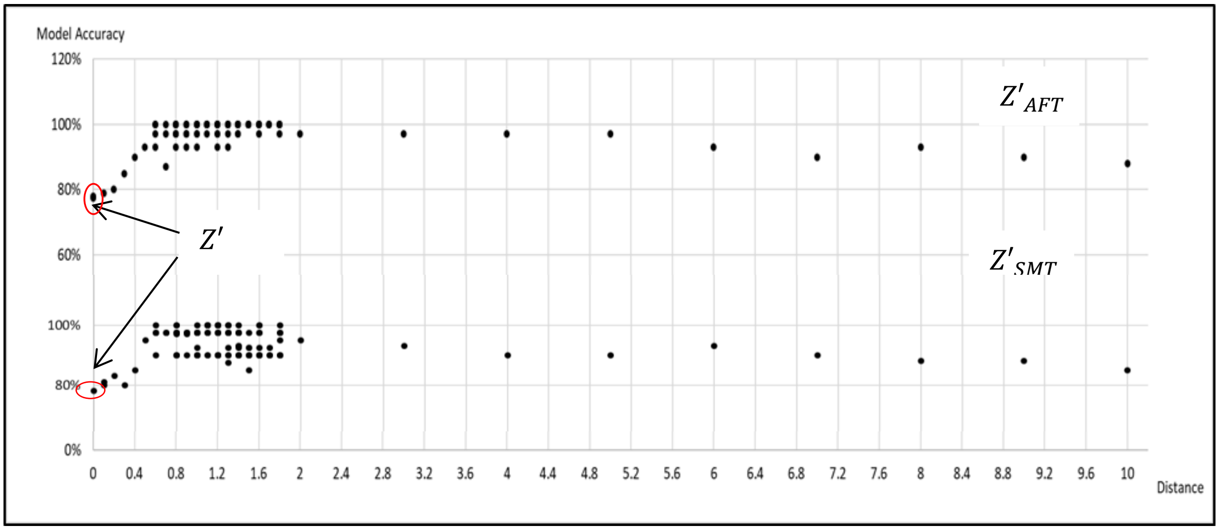


Figure 5-2 **Dependence of prediction accuracy on choice of scale.** Top graph: Z'_{AFT} model prediction accuracies; Bottom graph: Z'_{SMT} models accuracies. Model accuracy is calculated using the number of correct company classifications ($N_{correct}$/ N).

The Z'_{AFT} model accuracy (Figure 5-2, top graph) shows an upward trend in the accuracy rate from 78% to 90% for scale distances from zero to 0.4, respectively. When the distances further increase into the range of the 136 alternatives, the accuracy rate improves to between 90% to 95%. However, no trend within that range can be observed. For all scale distances beyond the 1.8 mark, the model accuracy remains indifferent to the particular scale choice for the AFT variable. The curve pattern suggests that the maximum increase in predictive accuracy of model (5.2) over model (4.6) is slightly above 10%. Based on this particular estimation sample, the critical point for the decision which scale to use, lies near a 0.5 distance for how to measure the AFT variable.

The prediction accuracy of the Z'_{SMT} model shows more variation compared to the Z_{AFT} model. For small scale distances of 0.1 to 0.4, the prediction accuracy increases slightly to about 78% to 83%. At the scale distance of 0.5 the accuracy rate jumps to 90%. For distances between 0.5 and 1.8, the prediction accuracy varies between 85% and 100%. It then falls back to a maximum of 90% for distances between 1.8 and 6.2, and further down to 80% as the distance keeps increasing. The overall trend suggests that the choice in the SMT variable distance is sensitive to the ability to forecast company failure. The optimal scale distance for the SMT variable thus lies within the range of between 0.5 and 1.8.

The data in Figure 5-2 show that both the Z'_{SMT} and Z'_{AFT} models outperform the Z' model for any choice of distances. However, a number of issues need be considered. Firstly, the forecasting performance is based on the estimation sample only. Do the results hold up with a different sample? (Section 5.2.1). Secondly, how do we explain the unexpected high and steady forecast accuracies of the Z'_{AFT} model for large distances? (Section 5.2.2). And thirdly, which particular scale distance should be chosen among the ranges? (Section 5.4).

5.2.1 Prediction success rates of Z'_{AFT} and Z'_{SMT} models

In Figure 5-3 I display the prediction accuracies of both models (5.1) and (5.2), and estimation and test samples. The distance of zero shows the success rate of the Z' model, i.e., without adding the AFT and SMT variable. In the Z'_{SMT} and Z'_{AFT} models the success rates for the estimation sample rates stay in the range of 90% and 95%. Whereas, success rates for the test sample out-performs the estimation sample success rates by reaching 100%. The Z'_{AFT_n} model better captures the test sample firms which means that the firms in the test sample are sensitive to the AFT variable.

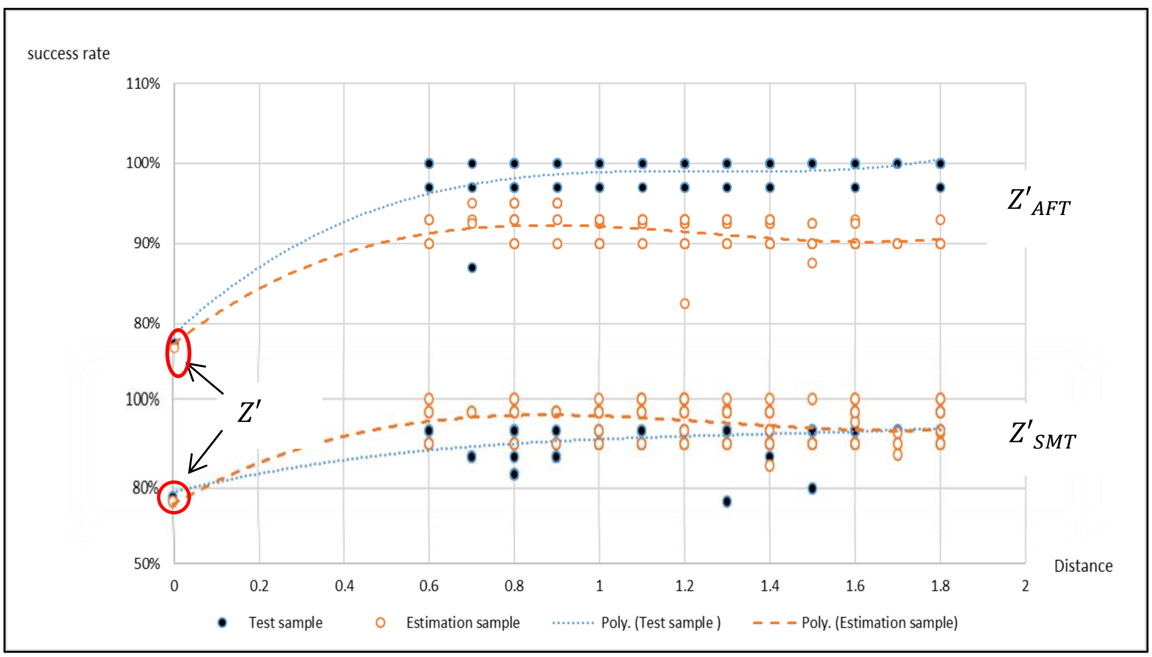


Figure 5-3 **Prediction success rates for both estimation and test sample** using 136 Z'_{SMT} and Z'_{AFT} models.

In the Z'_{SMT} model simulation, the estimation sample success rates outperform the test sample success rates. As stated previously, the Z'_{SMT} models' success rates are scattered which would be due to the high dominance of one type of error (either Type I or Type II). The estimation sample success rate ranges between 85% and 100% and the test sample range between 85% and 95%. All Z'_{AFT} and Z'_{SMT} models better forecast the sampled firms compared to the Z' model. The Z' model accuracy is between 75% and 80%.

In light of the test sample, the prediction accuracies remain robustly high independently of the distances in the AFT and SMT measurement scales

Figures 5-4 and 5-5 shows the Type I and Type II error spreads for both AFT and SMT simulations.

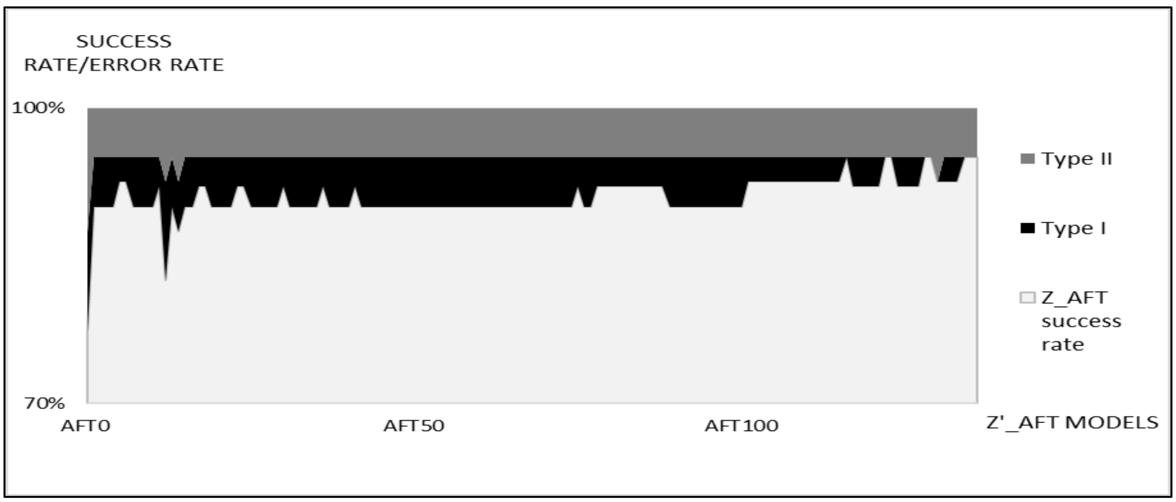


Figure 5-4 Success rates and the Type I and Type II error rates depicted for all the 136 Z'_{AFT} models using the estimation sample. AFT0 is the Z' model.

For the Z'_{AFT} model, Type I and Type II error across all 136 distance choices is relatively equally spread. The Type II error stays in the 5%-range consistently. Type I errors stay within the range of between 3% and 5%. Total misclassification is, on average, 10% which means 4 out of 40 firms are falsely categorised. At AFT0 (Z' model) the misclassifications Type I and Type II is approximately 10% and the total misclassification is 20%. All 136 Z'_{AFT} models outperform Z' model by fewer Type I and Type II errors and with better success rates.

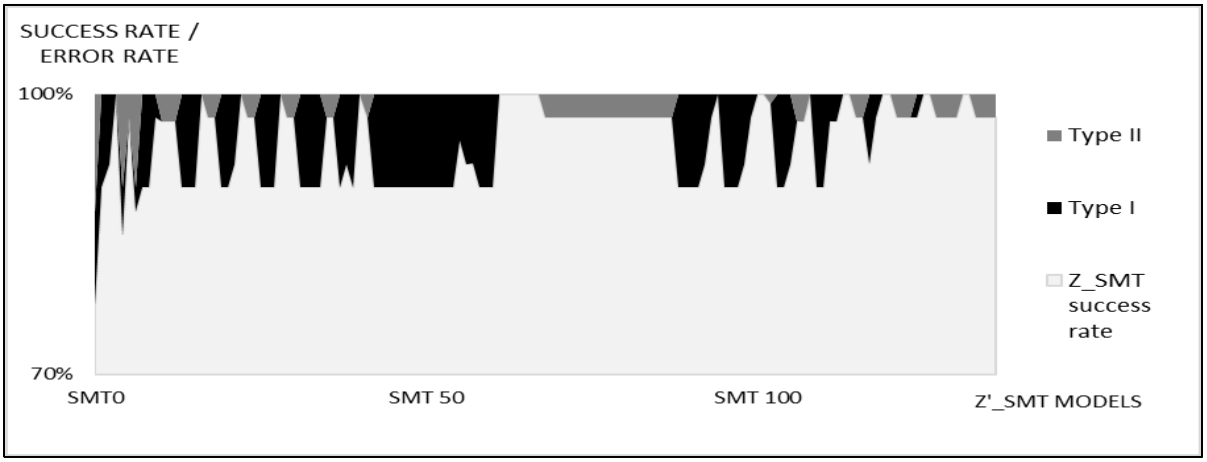


Figure 5-5 Success rate and Type I/II error rates depicted for all the 136 Z'_{SMT} models using the estimation sample. SMT0 is the Z' model.

Figure 5-5 shows a dominance of Type I error, i.e., most of the Z'_{SMT} models falsely identify non-failed companies as failed. Type I error is on average 10% (4 out of 40 firms). Type II error is within 0% to 3%. On average, the Z'_{SMT} model misclassifies up to 5 firms. Figure 5-4 and 5-5 shows that both Z'_{AFT} and Z'_{SMT} models outperform the Z' model. The success rates for all three models gives different rates of Type I and Type II errors. Considering that the simulation of the 2x136 models were run for the same estimation model, it is vital see whether or not the errors were made for the same companies as the Z' model did, or in fact, for different companies.

5.2.2 Individual firm misclassifications

The individual firm misclassifications for both Z'_{SMT} and Z'_{AFT} models are compared to the Z' model. When looking at the success rates it is important check the individual company success rates for all the 136 cases. This to check if the same companies got misclassified or different companies got misclassified. Since my simulation is run manually, one limitation is that I have a fixed allocation of firms into estimation and test samples. On the other hand, I would therefore expect that misclassifications will be committed on the same companies. This would particularly apply if the coefficients in the models stay relatively constant.

Table 5-2 displays the individual misclassifications for each company for all the 136 cases using my estimation sample.

Table 5-2 Summary of the individual firm misclassification for Z'_{SMT}, Z'_{AFT} and Z' models. Bold numbers represent firms misclassified as Type I errors and equally Type II errors are represented in normal font.

Model	Misclassified Firms	Same misclassification with Z' model	Different misclassification with Z' model	Total Number of Misclassifications
Z'	**2,14,16,18**,22,25,28,33,37	n/a	n/a	9
Z'_{SMT}	**14,16,17,20**,35	**14,16**	**17,20**,35	5
Z'_{AFT}	**18,20**,23,25	**18**,25	**20**,23	4

I expected the number of misclassified firms by the Z'_{SMT} and Z'_{ATF} models to be fewer than for the Z' model. The results in Table 5-2 show that both the Z'_{SMT} and Z'_{ATF} models make fewer errors but the misclassified firms include firms that were correctly classified by Z' model (firms 17, 20, 23 and 35). For the Z'_{SMT} model, the average total misclassification is 5 firms out of 40. This includes three firms that were already falsely classified by Z' model and three firms that were correctly classified by Z' model (firms 17, 20 and 35). However, the trade-off is that the Z'_{SMT} model correctly classifies seven more firm (firms 2, 18, 22, 25, 28, 33 and 37). All 136 cases for Z'_{SMT} model on average misclassify the five firms shown in Table 5-2. However, there are two cases where the firm misclassifications is different. In these two cases the firm misclassifies firms 14,16 as shown in Table 5-2 and firms 37,38,39, and 40 giving a total misclassification of 6. Apart from these two cases there were immaterially few other errors among the 136 distance choices observed in the misclassification (Appendix, Figure A-1).

The Z'_{AFT} model makes few mistakes as shown in Table 5-2. The Z'_{AFT} model on average falsely classifies 4 firms out of 40 in the estimation sample. These four firms include two firms (18 and 25) that were already falsely classified by the Z' model and two firms (20 and 23) that were correctly classified by the Z' model. As expected for all 136 Z'_{AFT} models, the same four firms were misclassified consistently

(Appendix, Figure A-2 shows that there were immaterially few other errors among the 136 distance choices observed in the misclassification).

This answers the issue of the curve pattern showed in Figure 5-2. The choice in the scale distance for AFT is indifferent between the values of 0.5 to 10, this due to the same companies being misclassified. Misclassification between 136 cases are indifferent to the added distance suggesting that the Z'_{AFT} successfully classifies 90% and the other 10% are insensitive to AFT variable thus the pattern on the curve steadily stays between 90% and 95%.

In summary, above discussion demonstrates the added value of analysing misclassification patterns on an individual firm level in order to understand model performance and model performance comparisons.

5.3 Coefficient behavior in the simulation

The success rate does not only depend on AFT and SMT variables but there are five other financial ratios that influence the change the in discriminant score. Therefore, it is vital to check how the parameters change given the change in the distance with which the qualitative variable are measured. For Equations (5.1) and (5.2) 136 linear regression models were run using MS Excel 2016. The parameters are denoted by $\beta_{n(1,..,6)}$ in the equations (5.1) and (5.2), and my interest lies in their distributional properties.

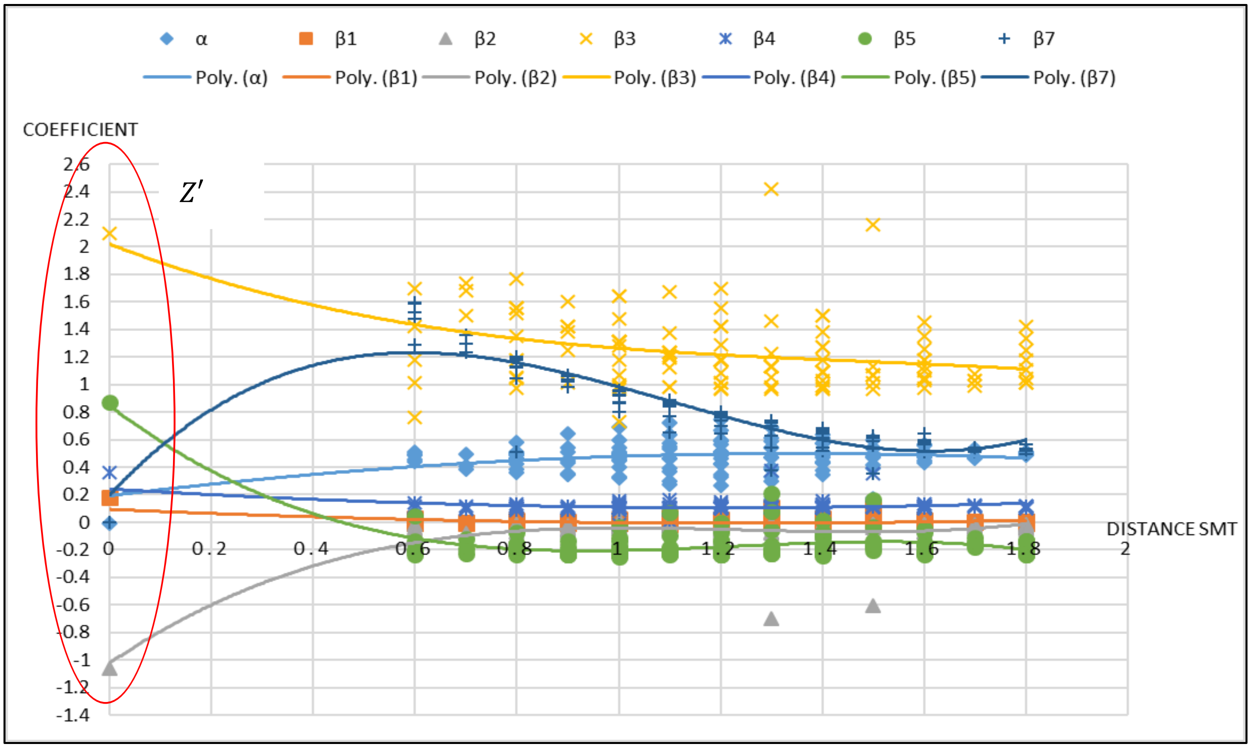


Figure 5-6 **Coefficient estimates vs scale distances for SMT.** Variations for the 136 linear regression models Z'_{SMT}.

Figure 5-6 shows the change in the coefficient estimates in relation to the change in the distance of the SMT measuring scale for 136 simulations. A zero distance represents the Z' model (4.6). In the Z' model, for example, the parameters estimates are highest for β_3 with 2.1, and β_5 at 0.9.

The trasition from a zero distance to a distance of 0.6 in the SMT variables shows a strong reaction to the other parameter estimates. That is, the addition of β_7 (SMT) has has reduced the weighting of β_5 (WCTA) from a positive to a negative contribution to the Z-score. The weighting of β_3 (EBITTA) has also decreased but remains the highest contribution to the discriminant score. β_2 (RETA) shows an upward tendency but the weightings approach zero from the negative. Both β_1 (STA) and β_4 (BVETL) does not show a strong reaction and stays near zero and 0.34 range consistently.

Notice that at distances between 0.6 to 1.8, apart from β_3 and β_7 all other weightings do not strongly change with an added distance for the SMT measurements. However, β_3 shows a decline in size towards larger SMT distance measurements. A possible explanation is that the coefficient estimates for β_3 vary the most, i.e., it has the highest average standard deviation from all variable parameters in the model with two of the 136 regression models estimating it at very high weightings of 2.4 (at distance 1.3) and 2.2 (at distance 1.5).

The coefficient estimates on SMT (β_7) show and interesting non-linear pattern. First, it is represented by a downward sloping curve which means as the scale distance increases, the coefficient weighting decreases. For distances larger than 1.6, the coefficient weighting seems to begin to increase.

The biggest drop is in β_5 (WCTA). In the Z_{est} model, β_5 was the second largest coefficient at 0.8. With the addition of the SMT variable the WCTA coefficient decreased to (-0.2) as if the variable now is almost 'invisible'.

Generally, the addition of the SMT variable has pressed many of the financial ratio multipliers in the Z-score to around zero, i.e., their contributions to classify firms as healthy have strongly decreased in size. Figure 5-6 also shows that the parameter estimates for β_1 , β_2 , β_4 and β_5 do not strongly fluctuate around their mean, i.e. the mean seems to be a robust estimate.

Figure 5-7 shows the equivalent analysis for the AFT variable as does Figure 5-6 for the SMT variable. β_3 and β_2 have the strongest reaction after the inclusion of the AFT variable. The transition from zero to 0.6 distance for β_3 shows an even steeper downward slope compared to that in Figure 5-6: the weighting decreases from 2.1 to nearly 0.4. The weighting of β_2 shows a similar upward trend when compared with Figure 4-6. β_1, β_4 and β_5 do not show a strong reaction and their coefficient weightings stay rather constant with any change in the scale distances of the AFT variable.

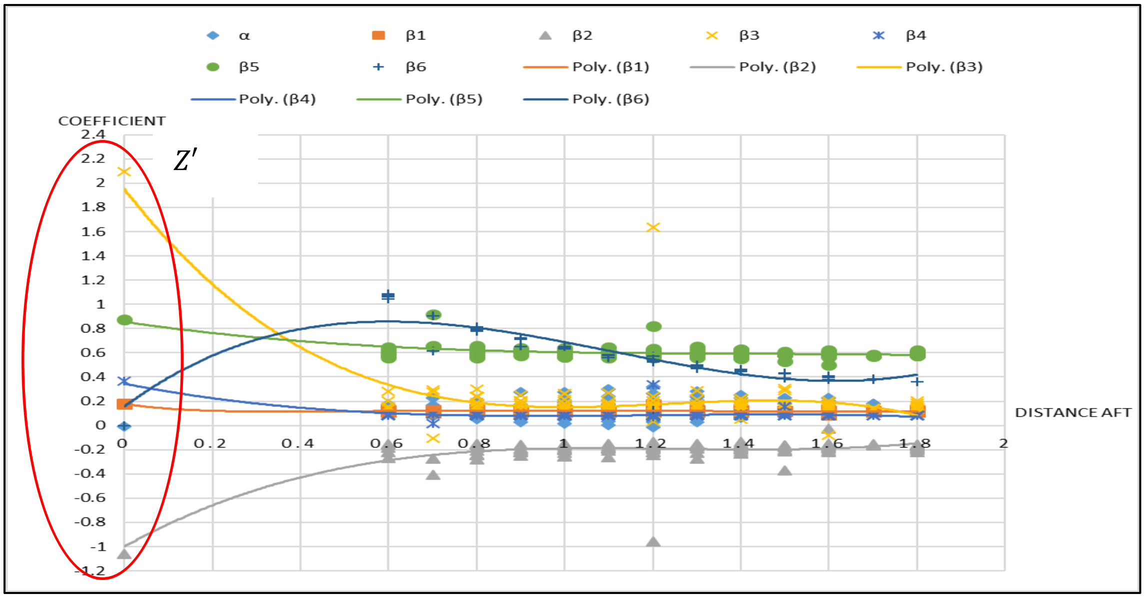


Figure 5-7 **Coefficient estimates vs scale distances for AFT.** Variations for the 136 linear regression models Z'_{AFT}.

The AFT coefficient β_6 pattern with changes in the scale distances is similar to that of the SMT coefficient shown in Figure 5-6. Unknown its values for distances between 0 and 0.6, the fitted polynomial may be misleading for it is possible that the values are very high for small distances. One of the different reactions from the inclusion of the SMT and AFT variables is observed in β_5: the effect of the SMT variable switches the sign of the WCTA variables; the effect of the AFT variable on β_5 does not show a dramatic change at all with the changes in the scale distances, however, it elevates β_5 to be the strongest multiplier on any of the five ratios.

Table 5-3 Comparisons of parameter estimates: Z' (model 4.6), Z'_{AFT} (model 5.2) and Z'_{SMT} (model 5.1). Means, standard deviations (in brackets) and size ranks from 136 different scale distances, as displayed in Figures 5-6 and 5-7.

Model	STA β_1	RETA β_2	EBITTA β_3	BVETL β_4	WCTA β_5	AFT β_6	SMT β_7
Z'	0.18	-1.06	2.10	0.40	0.90	n/a	n/a
Stdev	(0.07)	(043)	(0.93)	(0.18)	(0.4)		
Rank	(4)	(5)	(1)	(3)	(2)		
Z'_{SMT}	-0.01	-0.06	1.20	0.10	-0.17	n/a	0.8
Stdev	(0.04)	(0.25)	(0.49)	(0.10)	(0.25)		(0.08)
Rank	(6)	(4)	(1)	(3)	(5)		(2)
Z'_{AFT}	0.12	-0.2	0.19	0.09	0.60	0.56	n/a
Stdev	(0.06)	(0.37)	(0.79)	(0.15)	(0.31)	(0.1)	
Rank	(4)	(6)	(3)	(5)	(1)	(2)	

Table 5-3 shows the calculated average coefficients and standard deviations for all three models (4.6), (5.1) and (5.2). For the Z'_{SMT} model the means have declined in each parameter when compared to the Z' model. The parameters means for STA, RETA and WCTA are all negative and that would mean if STA is positive then it would be deducted from the discriminant score and vise versa. This sample dependency on the estimated parameter is concerning because one cannot expect that with increasing sales, given a constant total assets figure, a firm would be more likely to fail: but that is exactly what a negative β_1 represents. EBITTA and SMT have the largest contribution to the discriminant score. The average Z'_{AFT} coefficient estimates for WCTA, AFT and EBITTA dominantes the discriminant score as the other three coefficients are near zero.

As expected Table 5-3 shows for all the three models EBITTA to have the highest standard deviation. A high standard deviation means that there is a large variation in the coefficient estimate which in turn has the highest potential to be the reason for yielding differing failure prediction performances. Thus, when selecting an optimal distance scale measure for the two qualitative variables AFT and SMT, the change in coefficients and the the standard deviation needs to be considered. The smaller the standard deviation of the variable, the smaller the parameter bias of any one particular model (5.1) and (5.2) instance. Generally, recall that the Z'_{SMT} and Z'_{AFT} models outperform the Z' model with better forecast ability for both test and estimation sample.

5.4 Error term Normality and prediction success rates

The inclusion of ordinal variables with ratio variables in a linear regression model poses the risk of violating the assumption in error term Normality. AFT and SMT are both ordinal variables and the measures of these variables are discrete. Therefore the measures itself do not have a real meaning and for AFT and SMT I have high control in choosing the distances added to SMT and AFT variable. Whereas the five ratio variables hold real meaning in the values and I have no control over the measures of these variables. Both Z'_{SMT} and Z'_{AFT} models have a mixture of ratio variables and ordinal variables as independent variables and a dichotomous variable as the dependent variable. Use of dichotomous dependent variable and ratio independent variables are natural in the bankruptcy analysis. The mixture of ordinal and ratio variables in the analysis is seminal.

The error term Normality is tested on the Z', Z'_{AFT} and Z'_{SMT} models using Shapiro-Wilk test and Normal Q-Q plots. The Z' model's has a test significance of 0.259 (>0.05) and W statistic of 0.966. Normal Q-Q plot and test statistics suggest an approximate Normal distribution for the error terms. Figure 5-8 the distribution is approximately normally distributed and the Normal Q-Q plot shows the points are close to the Y=X line.

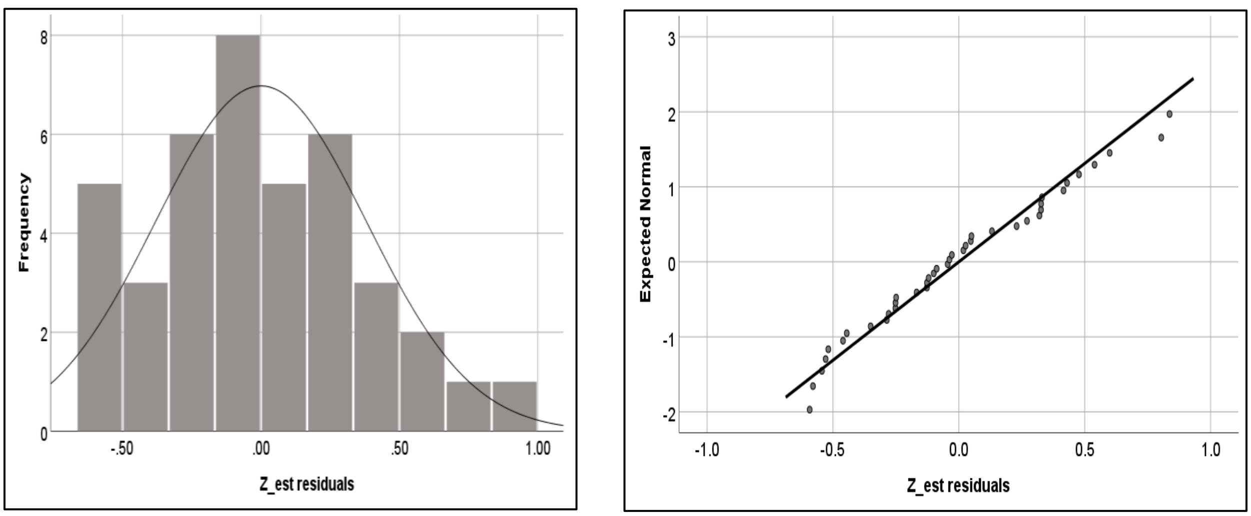


Figure 5-8 **Histogram and the Normal Q-Q plot for the Z' model error term (estimation sample).**

Results from testing the error term Normality for the 136 Z'_{SMT} and Z'_{AFT} models are shown in Figures 5-9 to 5-16.

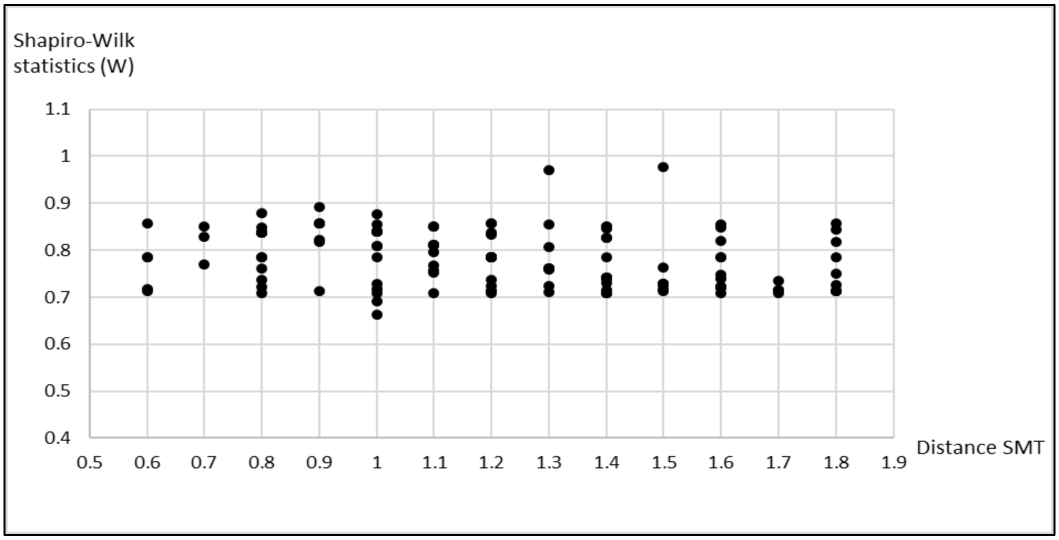


Figure 5-9 **Shapiro-Wilk W statistics against the SMT distance.**

Figure 5-9 shows 136 Z'_{SMT} model error term Normality statistics. The Shapiro-Wilk statistics are spread between 0.7 and 0.9 apart from two models that give a higher statistic of approximately 0.97 at SMT scale distances of 1.3 and 1.5. Thus the two preferred cases are SMT4 (at distance 1.5) and SMT6 (at distance 1.3) because their error terms approximate the Normal best. These two cases are analysed in Table 5-4 by success rates, misclassification rates and the model coefficient behaviour.

Table 5-4 **Analysis of favorable cases for SMT (SMT 4 and SMT6)**. The particular scale distances are as follows: SMT 4 (G1 0.6, G2 0, G3-0.9) and SMT 6 (G1 0.4, G2 0.2, G3 -0.9).

			Misclassifications					
	Shapiro-Wilk		Type I		Type II		Accuracy	
Case	W statistics	P-value	Estimation	Test	Estimation	Test	Estimation	Test
SMT4	0.977	0.573	5%	13%	10%	7%	85%	80%
SMT6	0.970	0.370	3%	13%	10%	10%	88%	70%

Panel B

	Coefficients					
	STA	RETA	EBITTA	BVETL	WCTA	SMT
Case	β1	β2	β3	β4	β5	β7
SMT4	0.10	-0.60	2.20	0.36	0.17	0.35
SD 4	0.06	0.39	0.77	0.16	0.41	0.09
SMT6	0.09	-0.12	0.66	0.14	0.21	0.45
SD 6	0.06	0.39	0.78	0.16	0.41	0.09

In Table 5-4 Panel A, comparison of the accuracy rates for both SMT4 and SMT6 shows that SMT6 makes more errors in the test sample 23% (both Type I and II) of the firms in test is misclassified (7 out of 30 firms). In contrast SMT4 misclassifies 20% of the firms in the test (6 out of 30). Both cases have a high dominance of Type I in the test sample and Type II error dominance in the estimation sample. SMT 4 has more errors in estimation sample compared to SMT6.

In Panel B the two cases have very similar coefficients for all the six variables and similar standard deviations. Highest contributor to the Z index is the EBITTA coefficient although it has a high standard deviation. The SMT coefficient is low since the distance added is high as seen in Figure 5-6. The choice of the optimal model would need to be compared to the Z' model. The Z' model has an accuracy of 78% for the estimation sample and 60% for the test sample. Therefore I would choose SMT4 as for both samples the SMT4 outperforms the Z' model.

The error coefficient behaviour for the two cases, SMT4 and SMT6, is visually observed using Normal Q-Q plots and histograms. The corresponding data are displayed in Figures 5-10 and 5-11. Both figures do agree that the error term is approximately normally distributed and the Normal Q-Q plots for both SMT4 and SMT6 show that most points lie close to the Y=X line (and when compared to that in Figure 5-8).

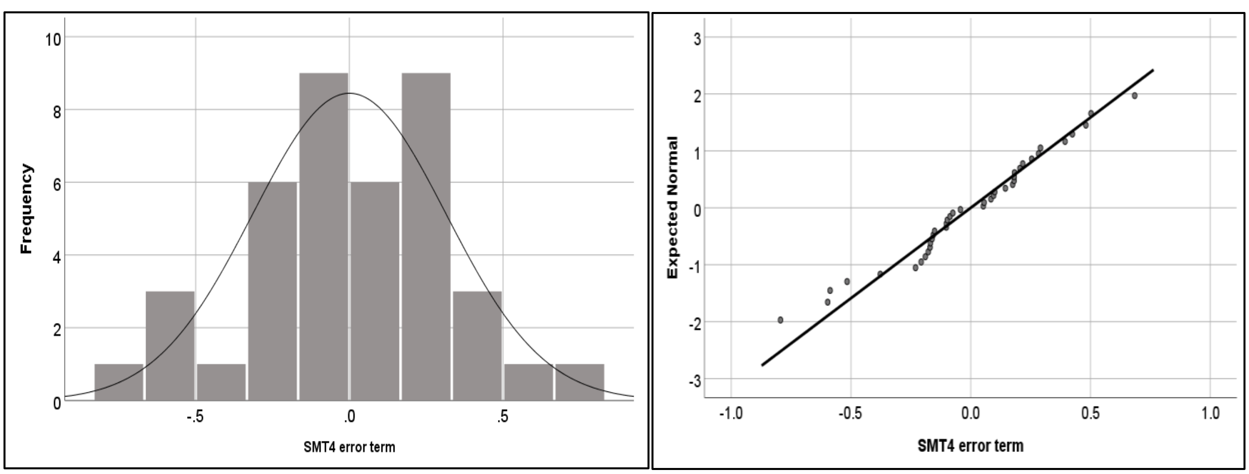


Figure 5-10 **Histogram and the Normal Q-Q plot observation for SMT4.**

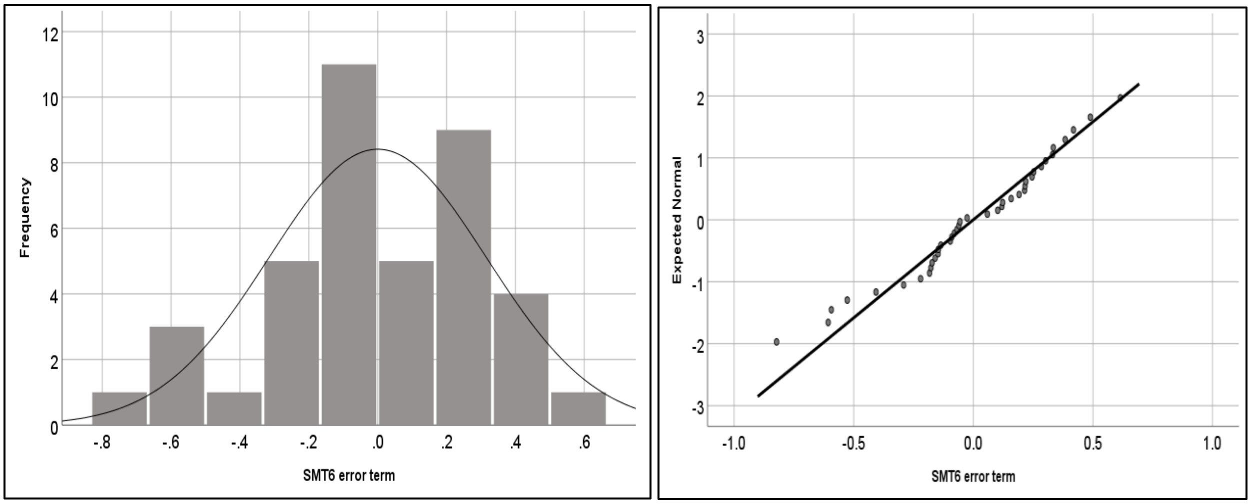


Figure 5-11 **Histogram and the Normal Q-Q plot observation for SMT6.**

In Table 5-3, the average Z'_{SMT} model, using the estimation sample, had a dominance of Type I errors. But in Table 5-4 we notice that both SMT4 and SMT6 have a dominance of Type II errors. Referring to Appendix Figure A-1 apart from SMT4 and SMT6 all the other models show a dominance for Type I errors. These two models are two odd cases but the misclassified firms in both cases are the same. Apart from these models all the other 134 Z'_{SMT} models violate the assumption of error term Normality, as shown next.

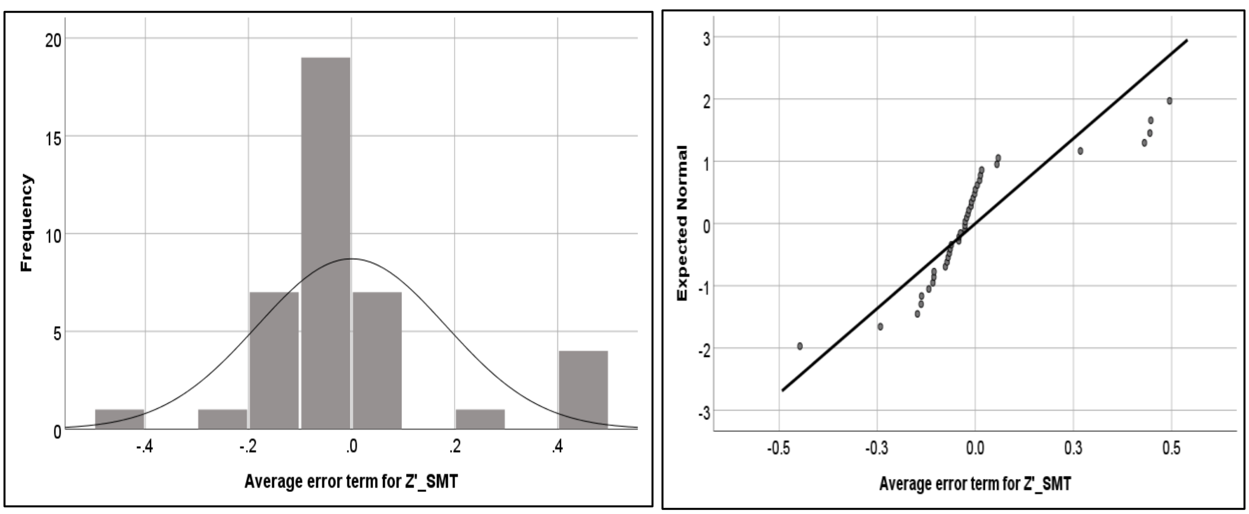


Figure 5-12 **Histogram and Normal Q-Q plot for the average Z'_{SMT} model error term**.

I now have averaged all the error terms from each of the 40 firms using the 136 different Z'_{SMT} model implementations to show how the average error term behaves across the 136 versions (cf. Figure 5-

12). The error terms are clearly not Normally distributed. The Shapiro-Wilk statistics is 0.791 which is below 0.94 and the significance is 0.00 (<0.05). Thus, the Shapiro-Wilk statistic also shows that the average Z'_{SMT} model error term is not Normally distributed. In the histogram the standard deviation for the data is 0.18 which yields a very narrow curve compared to the fitted standard normal curve (mean = 0 and SD =1).

Therefore I choose the SMT4 scale distance settings as the optimal measure to value the SMT variable since it fulfils the normality assumption. The estimated model using the estimation sample yields:

$$Z'_{SMT_4} = 0.57 + 0.1STA - 0.6RETA + 2.2EBITTA + 0.36BVETL + 0.17WCTA + 0.35SMT \quad (5.4)$$

What follows is an identical analysis for the AFT variable. The same tests were run for all 136 Z'_{AFT} models. A similar pattern is observed for the Z'_{AFT} error terms. Figure 5-13 shows the Shapiro-Wilk statistics in relation to the scale distances of the AFT variable.

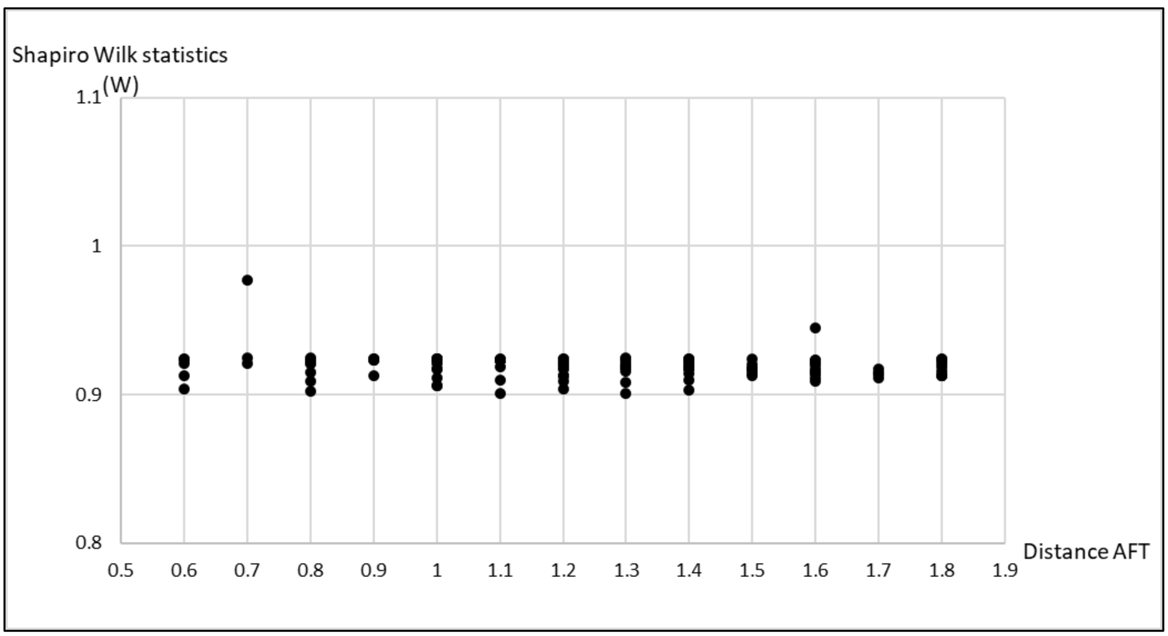


Figure 5-13 **Shapiro-Wilk statistics (W) against the AFT distances for the Z'_{AFT} model.**

Compared to Figure 5-9, the error term for Z'_{AFT} models are within 0.9 and 0.94 apart from one model that is close to 0.97. In the individual model analysis there were two models that have W statistics above 0.94 and a significance above 0.05: one is at the scale distance of 0.7 (AFT41) and the second model is at a scale distance of 1.6 (AFT97). First, I thus analyse these two candidate cases AFT41 and AFT97 in Table 5-5 using success rates, misclassification rates and the model coefficient behaviour.

In Table 5-5 Panel A, I compare the Shapiro-Wilk statistics between the AFT97 and AFT41 models using the estimation sample; and I also compare the failure prediction accuracy rates for both estimation and test samples. AFT97 has a Shapiro-Wilk significance of 0.053 and a high failure prediction accuracy in both estimation and test samples. AFT41 has significance of 0.576 but a lower accuracy in both estimation and test samples. Thus both models seem to meet the error term Normality assumption barely.

AFT97 predicts better in the test sample compared to in the estimation sample. AFT41 predicts better in estimation sample but predicts less in the test sample. Overall AFT97 outperforms AFT41 in its predictive performance.

Table 5-5 **Analysis of favorable AFT case (AFT41 and AFT97).** The particular scale distances are as follows:AFT41 (G1 0.4, G2 0.2, G3 -0.3) and AFT97 (G10.8, G2-0.2, G3-0.8).

			Misclassifications					
	Shapiro-Wilk		Type I		Type II		Accuracy	
Case	W statistics	P-value	Estimation	Test	Estimation	Test	Estimation	Test
AFT41	0.977	0.576	3%	3%	5%	10%	93%	87%
AFT97	0.923	0.053	5%	0%	5%	0%	90%	100%

Panel B

	Coefficients					
	STA	RETA	EBITTA	BVETL	WCTA	AFT
Case	β1	β2	β3	β4	β5	β6
AFT41	0.11	-0.41	-0.11	0.02	0.91	0.62
SD41	0.06	0.36	0.87	0.16	0.32	0.13
AFT97	0.09	-0.02	-0.08	0.09	0.50	0.38
SD97	0.06	0.39	0.84	0.15	0.32	0.08

Panel B in Table 5-5 shows the coefficients for the two models using AFT41 and AFT97. In AFT41 the coefficient weighting is 0.62 and almost double of that in AFT97 at 0.38. This may be because the maximum scale distance in AFT41 is 0.7 which is about half of the scale distance in AFT97 which is 1.6. All the other five coefficients are smaller for AFT97 compared to AFT41. And both AFT41 and AFT97 models outperform the Z' model in their predictive accuracy. Hence, there are no reasons which would make either of the two models AFT41 and AFT97 a better choice.

Normal Q-Q plots and histogram for both AFT41 and AFT97 are shown in Figures 5-14 and 5-15. AFT41 error terms fit the Normal distribution better compared to AFT97. AFT97 has many more error terms near zero.

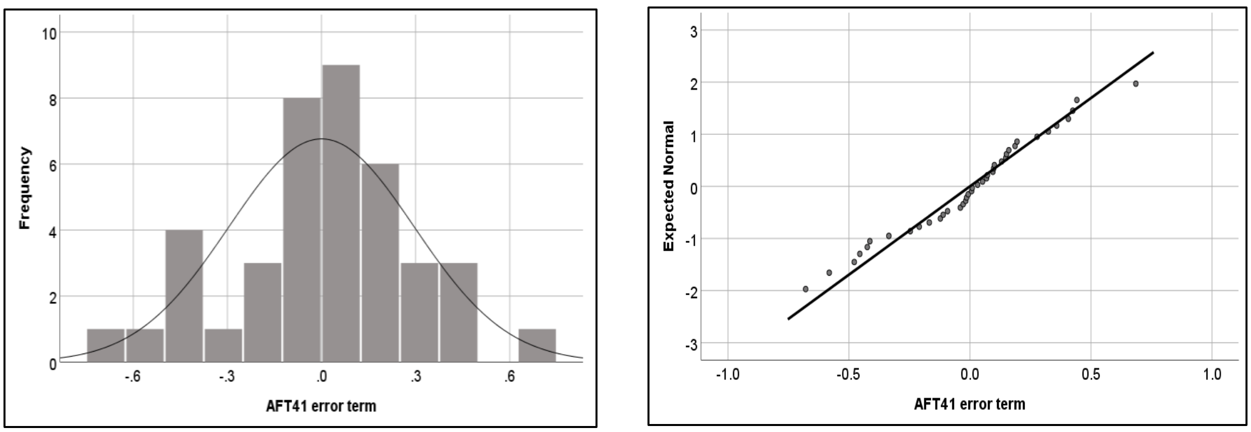


Figure 5-14 **Histogram and the Normal Q-Q plot for AFT41.**

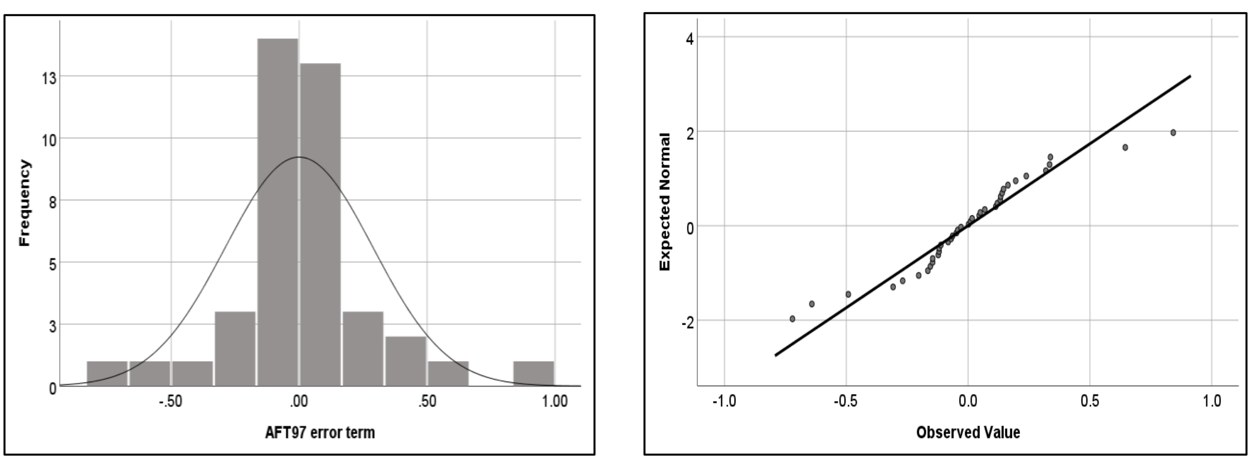
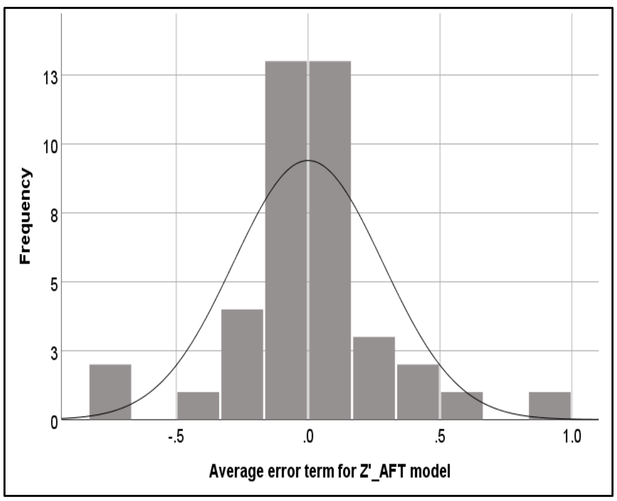


Figure 5-15 **Histogram and the Normal Q-Q plot for AFT97.**

The AFT97 model has a lower Shapiro-Wilk significance compared to AFT41 which corresponds to the data shown in the figures. Both Normal Q-Q plots in Figures 5-14 and 5-15 fit the Y=X line equally badly as the two SMT models do (cf. Figures 5-10 and 5-11).

Finally, I investigate the general error term behaviour for the AFT models. Figure 5-16 shows the average error distribution for the estimation sample across the 136 Z'_{AFT} models. The patterns shown in the data are more similar to the AFT97 than the AFT41 model.



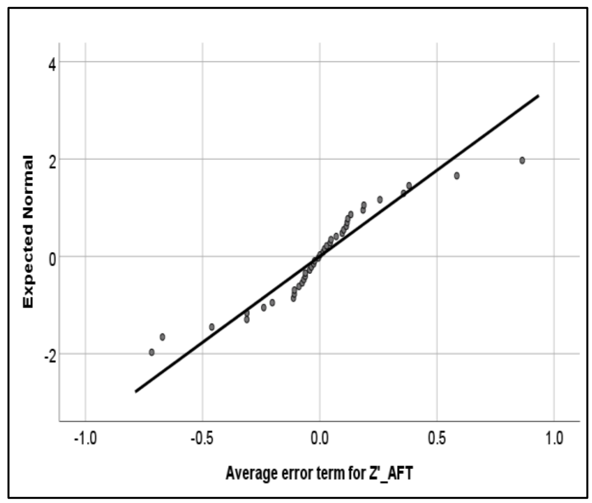


Figure 5-16 **Histogram and Normal Q-Q plot for the average Z'_{AFT} model error term.**

The average Z'_{AFT} error term Shapiro-Wilk W statistic is 0.926 with a significance of 0.012 (<0.05). As shown in Figure 5-13 the Shapiro Wilk statistics for the Z'_{AFT} error term is clusters around 0.9. Hence the average Z'_{AFT} error term spread (Figure 5-16) is similar to that of Figure 5-15.

From the two choices AFT41 and AFT97 analysed, I choose AFT97 because it better predicts company failure in my sample.

The data above seem to provide for a paradox. Adding the SMT (AFT) variable to the Z' model improves (worsens) the error term approximation to Normality but has a lower than AFT (higher than SMT) increase in predictive accuracy compared to the Z' model.

Irrespective of this interesting model behaviour, the two measures that I have chosen to value AFT and SMT are AFT97 and SMT4, respectively. The corresponding models are as follows (Equation 5.4 is reprinted here for convenience):

$$Z'_{SMT_4} = 0.57 + 0.1STA - 0.6RETA + 2.2EBITTA + 0.36BVETL + 0.17WCTA + 0.35SMT \quad (5.4)$$

$$Z'_{AFT_{97}} = 0.22 + 0.09STA - 0.02RETA - 0.08EBITTA + 0.09BVETL + 0.50WCTA + 0.38AFT \quad (5.5)$$

The corresponding scale distances are:

- for the SMT variable: G1=0.6, G2=0.0 and G3=-0.9; and
- For the AFT variable: G1=0.8, G2=0.2 and G3=-0.8.

These two models will be used to in the following Section 5.5 and Chapter 6. In Section 5.5 I compare them to a 7-variable Z-score model that includes both the AFT and SMT variables, and in Chapter 6, I use all three models to check to what degree these can predict the Carillion PLC company failure.

5.5 The 7-variable Z-score model using AFT and SMT variables

All procedures (error term Normality, Shapiro-Wilk, Normal Q-Q plots and predictive accuracy) from the previous section were run for the $Z'_{AFT,SMT}$ model (5.3) which I reprint for convenience:

$$Z'_{AFT,SMT} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_6 AFT + \beta_7 SMT \quad (5.3)$$

In order to find a best scale distance combination for the AFT and SMT variables, I would need to search in a 136*136 matrix of possible permutations; however, due to the robustness in the prediction performance I choose the 136 diagonal cases to test the performance. Clearly, searching the space of all reasonably possible scale distance combinations is outside the scope if this work, but may provide for an interesting future research project. The individual misclassification for all the 136 cases are run (cf. Appendix Table A-3) and the misclassification results are shown in Table 5-6.

Table 5-6 **Summary of the individual firm misclassifications for $Z'_{AFT,SMT}$.** Misclassification are compared against the Z' model. The observations are for the estimation sample. Bold numbers represent firms misclassified as Type I errors and equally Type II errors are represented in normal font. Firms with a '*' sign are misclassified in only 3 out of 136 cases (refer to Figure A-3 in Appendix).

Model	Misclassified Firms	Same misclassification with the Z' model	Different misclassification with the Z' model	Total Number of Misclassifications
Z'	**2,14,16,18**,22,25,28,33,37	n/a	n/a	9
$Z'_{AFT,SMT}$	**16,20**,23*,25*,35	**16**,25*	**20**,23*,35	5

Table 5-6 shows all the classification errors made by any of the 136 renditions of the $Z'_{AFT,SMT}$ model. The generic $Z'_{AFT,SMT}$ misclassifies a maximum of five firms, where any of the 136 individual implementations misclassifies a maximum of three firms and a minimum of one (c.f. Figure A-3 in Appendix). There are two firms (16, 25) that are being misclassified by the Z' model too. There are three firms (20, 23, and 35) that were misclassified by the $Z'_{AFT,SMT}$ model, but not by the Z' model. Recall that in Table 5-2 the Z'_{SMT} and Z'_{AFT} models both made a classification error on firm 20, and that Z'_{AFT} falsely classified firms 23 and 25, and Z'_{SMT} falsely classified firm 35. Thus, these 'problematic' firms were signalled earlier either by the application of the (4.6), (5.1) or (5.2) models. Also the trade-off in using this model is that it correctly predicts firms 2,14,17,18,22,28,33 and 37 which were falsely classified by the Z' model. The $Z'_{AFT,SMT}$ clearly outperforms the Z' model in its prediction ability.

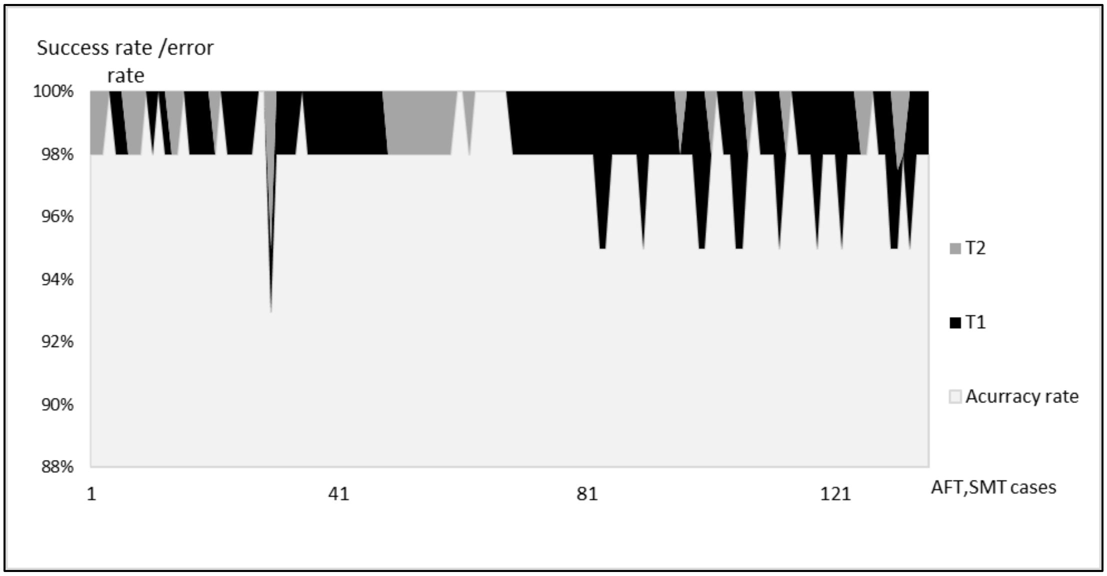


Figure 5-17 **Success rates and Type I/II error rates for $Z'_{AFT,SMT}$ model.**

The $Z'_{AFT,SMT}$ success rates across the 136 models are mainly around 93% to 100%. Type I errors dominate. Firm 20 has been falsely classified in many cases here (cf. Figure A-3 in the Appendix) and also by the Z'_{AFT} and Z'_{SMT} models (cf. Table 5-2). The maximum risk of this model is 7% (4 out of 40 firms) which gives it a high robustness for the choices between the 136 combinations of AFT and SMT scale distances tested.

I have run the same analysis for $Z'_{AFT,SMT}$ model as in Section 5.3 to observe the coefficient behaviour of the model. The average of each coefficients are calculated for the 136 cases and the following Figure 5-18 shows the average coefficient weighting for Equations (5.1), (5.2) and (5.3) and compare these to model (4.6). To clarify, I use the respective scale distances for the SMT model (5.1) and AFT model (5.2), which is found on the diagonal of the 136x136 matrix used for model (5.3).

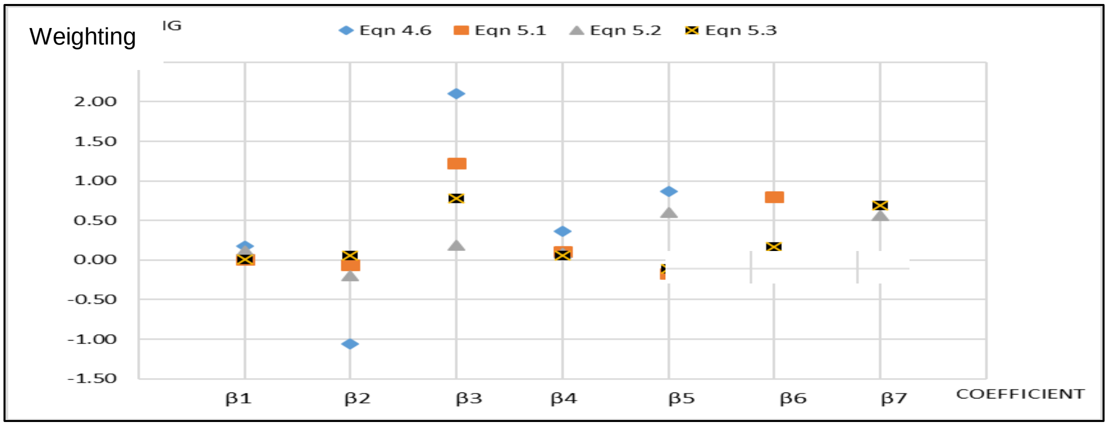


Figure 5-18 **Average coefficients weighting for all models (4.6), (5.1), (5.2) and (5.3).** (Table 5-7)

In Figure 5-18, β_3 shows more variation across the different models as compared to all the other coefficients. It is also the largest of all weightings, on average. As alleged in Section 5.3 the EBITTA ratio has a high standard deviation and it is not surprising to see the change in the parameter variation.

Table 5-7 **Coefficient weighting comparison for all four models.** For Equations (5.1), (5.2) and (5.3) coefficients show are the averages from the simulations.

		STA	RETA	EBITTA	BVETL	WCTA	AFT	SMT
	α	β1	β2	β3	β4	β5	β6	β7
Eqn 4.6	-0.01	0.18	-1.06	2.10	0.36	0.87	n/a	n/a
Eqn 5.1	0.48	0.01	-0.06	1.22	0.11	-0.17	0.80	n/a
Eqn 5.2	0.15	0.12	-0.20	0.19	0.09	0.60	n/a	0.56
Eqn 5.3	0.45	0.01	0.06	0.78	0.06	-0.11	0.17	0.69

Table 5-7 shows a detailed breakdown of Figure 5-18. Generally, the average parameter estimates have changed for the $Z'_{AFT,SMT}$ model (5.3), as they did for the Z'_{AFT} and Z'_{SMT} models earlier. The EBITTA and SMT coefficients are large in comparison with all the other variables which are near zero which suggests that EBITTA and SMT will dominate the discriminant scores. Therefore, consideration must be given to the behaviour of the coefficients when choosing the combined measuring scales for the AFT and SMT variables.

So now I choose the two points for the combined AFT and SMT model by considering the error term normality and the predictive accuracy of the particular $Z'_{AFT,SMT}$ models.

5.5.1 Error term Normality

I run the chosen 136 cases for the $Z'_{AFT,SMT}$ model, and in addition, I also choose the two optimal scale distances found earlier in Section 5.4 for the 6-variable models (AFT97, SMT4) to check the error term Normality. I complete my analysis with the Shapiro-Wilk statistic and Normal Q-Q plots that are observed for the 137 cases.

The Shapiro-Wilk test showed worsened results for the Normality test of the $Z'_{AFT,SMT}$ error terms. The W-statistic yielded 0.86 which is lower than 0.94 and the significance is 0.00 which suggest that the error term is not Normally distributed.

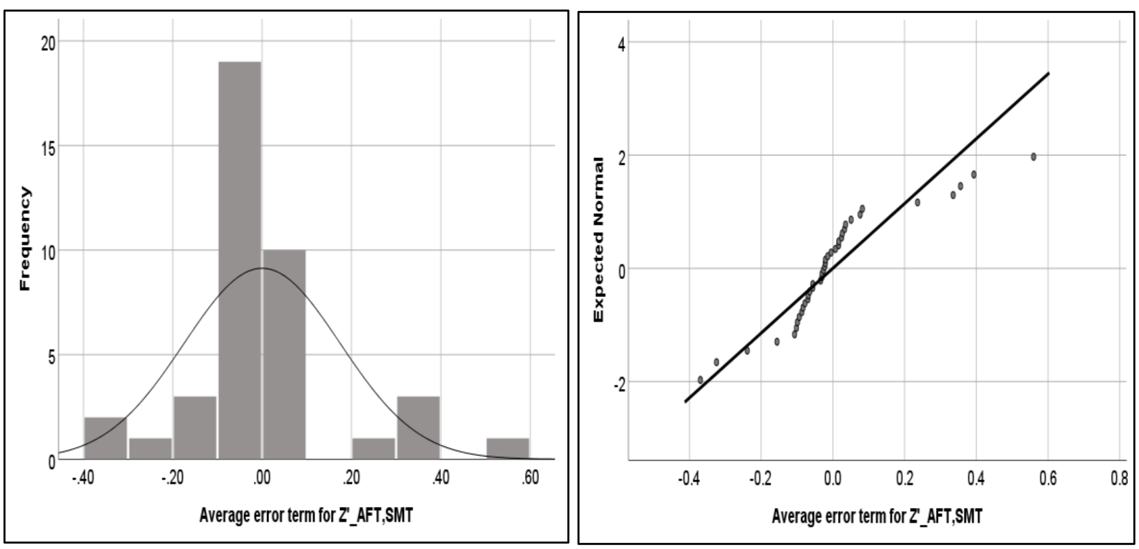


Figure 5-19 **Histogram and the Normal Q-Q plot for the $Z'_{AFT,SMT}$ model.**

Figure 5-16 shows that there is in fact a high deviation of the average error term from the Normal distribution. In the histogram we observe that the high number of points clusters around zero and also there are some long tails. This figure can be compared with Figure 5-12 where the SMT simulation gave a similar outcome. Although the $Z'_{AFT,SMT}$ outperforms the Z' model by predictive accuracy the non-fulfilment of the error term Normality may concern.

5.5.2 Optimal value that satisfy both error term Normality and predictive performance

The Shapiro-Wilk analysis with the highest W and associated significance was obtained for the combination (AFT131,SMT6). I choose to compare that model with the (AFT97,SMT4) combination, as based on my analysis in Section 5.4, to choose the optimal model that would best satisfy both error term Normality as well as improve the predictive performance in my sample. The error term Normality statistics and predictive performances for both estimation and test samples are shown in Table 5-8.

Table 5-8 **Analysis of potentially best performing cases.**

Panel A

			Misclassifications					
	Shapiro-Wilk		Type I		Type II		Accuracy	
Case	W statistics	P-value	Estimation	Test	Estimation	Test	Estimation	Test
AFT131,SMT6	0.923	0.100	5%	0%	0%	0%	95%	100%
AFT97,SMT41	0.947	0.058	3%	0%	0%	0%	98%	100%

Panel B

	Coefficients						
	STA	RETA	EBITTA	BVETL	WCTA	AFT	SMT
Case	β1	β2	β3	β4	β5	β6	β7
AFT131,SMT6	0.093	-0.117	0.664	0.140	0.211	0.454	0.269
AFT97,SMT41	0.069	0.023	0.382	0.144	0.141	0.302	0.237

The error term normality statistics for (AFT97,SMT4) are W=0.947 with a significance greater than 0.05. Similar statistics are found for the (AFT131,SMT6) combination. The error term approximation to Normality thus will not provide for a clear preference.

According to the predictive performance of both models, the (AFT97,SMT4) model outperforms the other case. I therefore use the (AFT97,SMT4) model with corresponding scale distances as my optimal case to represent the $Z'_{AFT,SMT}$ score. This model, according to Panel B of Table 5-8 has positive coefficient signs on all of the 7 variables.

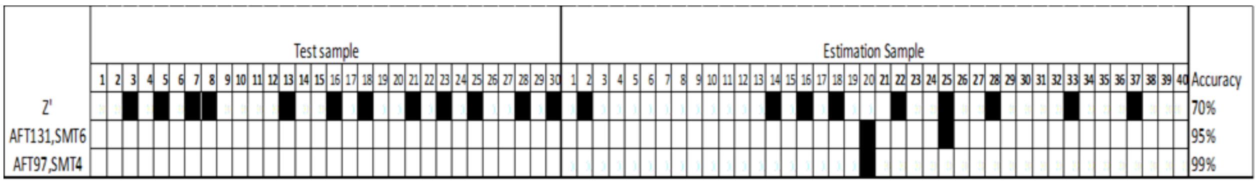


Figure 5-20 **Individual misclassifications for all three chosen models.**

In Figure 5-20, I compare the classification patterns of the three models at firm-level. As noted earlier, firm 20 has been falsely classified by both extended models and not by the Z' model. This suggest that firm 20 is not sensitive to the AFT and SMT variables. However, the trade off in using the $Z'_{AFT97,SMT4}$ model is that it predicts much better in general: 99% (98.6 rounder to 1dp) of all firms in my entire sample have been correctly classified 1 year ahead of potential failure.

From the above analysis it is clear that I choose model $Z'_{AFT97,SMT4}$ to now analyse the case of Carillion PLC. The final 7-variables model is as follows:

$$Z'_{AFT,SMT} = 0.28 + 0.07STA + 0.02RETA + 0.38EBITTA + 0.14BVETL + 0.14WCTA + 0.30AFT + 0.24SMT \quad (5.6)$$

5.6 Discussion

Any model using any combination of qualitative variables and associated scale differences in the range between 0.6 and 1.8 outperform Z’ model (4.6). The predictive accuracy of the 6- and 7-variable models showed a robust and steady performance for all reasonable scale distances. However, there was a negative correlation between error term Normality and the predictive performance. The four models analysed throughout this chapter are reprinted below:

$$Z' = -0.01 + 0.18STA - 1.06RETA + 2.1EBITTA + 0.4BVETL + 0.9WCTA \quad (4.6)$$

$$Z'_{SMT_4} = 0.57 + 0.1STA - 0.6RETA + 2.2EBITTA + 0.36BVETL + 0.17WCTA + 0.35SMT \quad (5.4)$$

$$Z'_{AFT_{97}} = 0.22 + 0.09STA - 0.02RETA - 0.08EBITTA + 0.09BVETL + 0.50WCTA + 0.38AFT \quad (5.5)$$

$$Z'_{AFT,SMT} = 0.28 + 0.07STA + 0.02RETA + 0.38EBITTA + 0.14BVETL + 0.14WCTA + 0.30AFT + 0.24SMT \quad (5.6)$$

The corresponding classification patterns on an individual firm-level are shown in Figure 5-21.

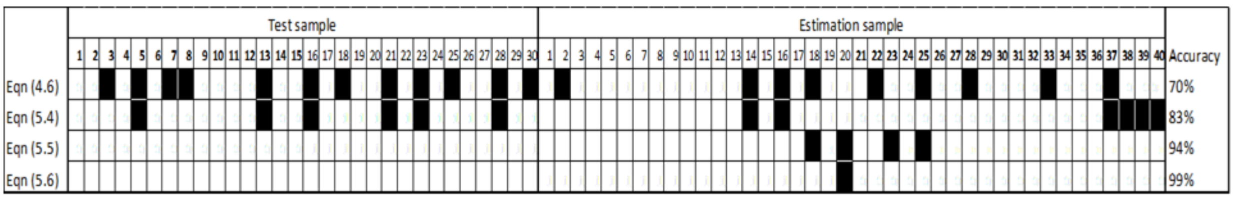


Figure 5-21 **Individual misclassifications for all 70 firms in my sample.**

Models (5.5) and (5.6) both correctly predict all the firms in my test sample. In the estimation model (5.6) produces the best performance with only one false misclassification. The choice of best performing model thus is clearly identified. However, I have found a paradox result in a negative correlation between the error term approximation to Normality and the predictive performance. That is, the worse the error term approximates the Normal distribution (low Shapiro-Wilk p-values), the larger the chance for an omitted variables problem, which would make a model less useful to predict. However, this effect is reversed as shown in Figure 5-22.

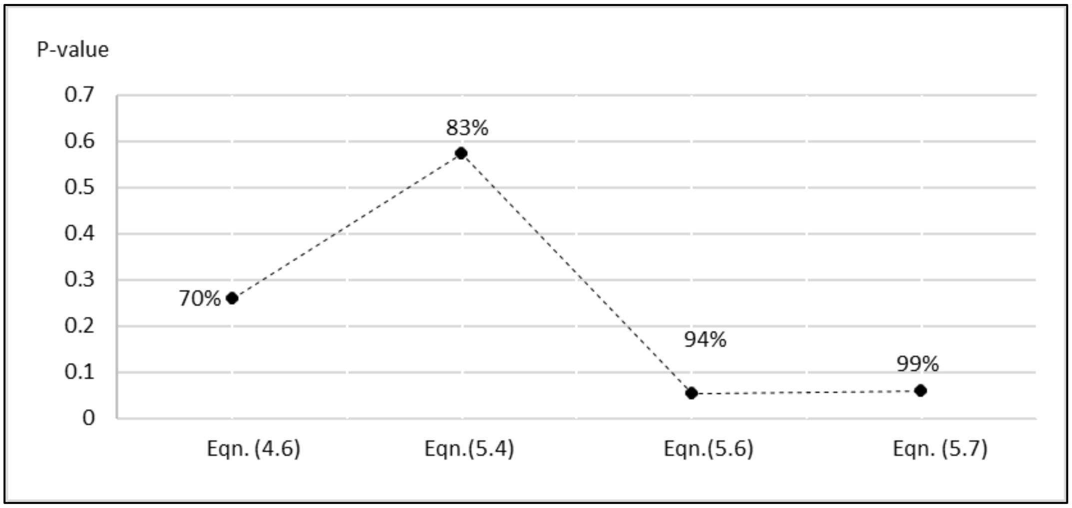


Figure 5-22 **Shapiro-Wilk significance (p-value) and the prediction success rates.**

The figure shows a negative correlation between the Shapiro-Wilk significance and the success rates for each model.

Chapter 6 Carillion PLC analysis

The construction sector in the UK was chosen due to the timely application of Carillion PLC which collapsed in early 2018. Carillion PLC was one of the largest construction companies in the UK. Carillion PLC was incorporated in 1999 at which time it was headquartered in Wolverhampton, UK. When Carillion filed for liquidation on 15th January, 2018, the company directors and the auditors (Internal auditor Deloitte and external auditor KPMG) were held liable for accounts and false assurance given to the investors. In the year 2016/2017 the company has collapsed, five directors were terminated, and during the liquidation, KPMG suspended the lead partner on Carillion audit over issues related to the documentation provided to the Financial Reporting Council. The Financial Reporting Council regulates accountants, auditors and actuaries by prescribing UK's corporate governance system.

As motivated in the introduction, I investigate whether the Carillion PLC failure could have been predicted. Financial data for Carillion from 1999 till 2016 are collected and analysed using all four models Z' ,Z'_{SMT4}, Z'_{AFT97} and $Z'_{AFT97,SMT4}$ models. Figure 6-1 shows the trend analysis for Carillion's Z-scores.

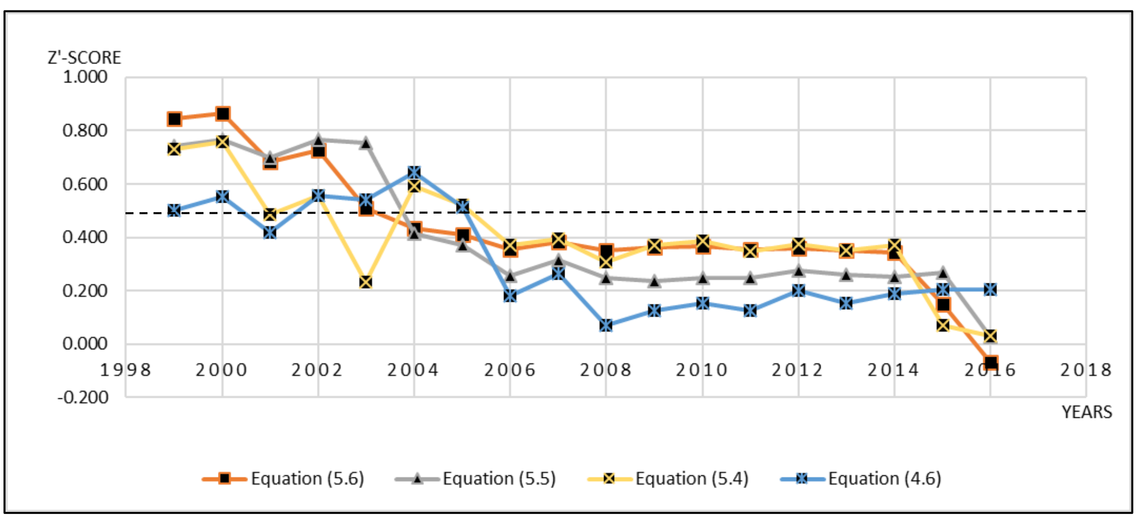


Figure 6-1 **Z-score trend analysis for Carillion using four models.** In all cases, the cut-off value is 0.5.

In Figure 6-1 it is apparent that Z'_{SMT4}, Z'_{AFT97} and $Z'_{AFT97,SMT4}$ signal financial distress already in 2015 and 2016 as a consequence of a higher director turnover in this period. The Z'_{AFT97} model (grey line) does capture the decline one year later in 2015 than the Z'_{SMT4} and $Z'_{AFT97,SMT4}$ models because the director changes are captured by the SMT variable. It is interesting to notice that from 2005 to 2016 all Z-scores for all the models fell to below 0.5 which is the cut-off point of the models.

Table 6-1 **Breakdown of the final four model coefficient weightings.**

		Coefficients						
		STA	RETA	EBITTA	BVETL	WCTA	AFT	SMT
	α	β1	β2	β3	β4	β5	β6	β7
Eqn(4.6)	-0.01	0.18	-1.06	2.1	0.4	0.9	n/a	n/a
Eqn(5.4)	0.57	0.1	-0.6	2.2	0.36	0.17	n/a	0.35
Eqn(5.5)	0.22	0.09	-0.02	-0.09	0.09	0.5	0.38	n/a
Eqn(5.6)	0.28	0.07	0.02	0.38	0.14	0.14	0.3	0.24

In Equation (5.6) the coefficients for variables EBITTA, AFT and SMT are large compared to the other variables. Depending on the size of the financial ratio, we see a sudden drop in the Z-score in 2015 where the SMT was high. Note that Equation 5.5 only drops in 2015 when KPMG, their main audit partner, was dismissed.

Carillion PLC's Z-score in each of the four models are different and all are below the cut-off point. Then a question arises at what point would one deem the company a failing one? Even just using the Z' model we can clearly see a downward trend in the financials but the drops in 2008 can be due to the global financial crisis and drop in 2011 was due to the recession in UK construction sector. Using the Z' model (4.6), one could also argue that the upward trend starting from 2014 suggest that the company is recovering from lower performance levels.

In summary, there should not have been a surprise that Carillion PLC failed. A simple Z-score analysis that any accounting and finance literate investor or an industry overseeing agency could have made, would flag this company at a higher or high risk of failure over an extended period of time (cf. Figure 6-1).

However, a Z-score that includes only financial ratios would not have signaled high company distress during the past 3 years. This could only have been detected by adding variables to the model sensitive to other failure modes. My bankruptcy analysis in Chapters 1 to 5 substantiate these claims.

Chapter 7 Presentation of bankruptcy analysis in accounting education

The Altman Z-score (1968) is the most popular bankruptcy prediction model that appears in accounting textbooks. The contexts given within which the Z-score is presented can lead to mis-applications for the lack of necessary information provided which any accounting student must be made aware of before they mistakenly evaluate firm performance with the original Z-score or re-estimations thereof. Necessary information includes:

- Variable selection: Companies fail for lots of different reasons. Have the readers been informed about that the variables in the model may not capture all of the failure modes?
- Model Estimation – parameter coefficients: Parameter coefficients shown are not universally representative. They depend on the estimation sample. Have the readers been made aware of the time-sensitivity of those coefficients? Most users will apply Z-scores now and would have to re-estimate the coefficients every time new financial information has been released.
- Model Estimation – sample: There are many ways how to obtain, slice and dice a sample that will determine variable selection, parameter estimation and model testing. Have the readers been made aware that the presentation of the model depends on the underlying sample?
- Model Usefulness: "All models are wrong, but some are useful" (G. Box). Have the readers been informed about the model's predictive ability subject to any assumptions made?

I searched accounting text books to explore and to understand how the Z-score model is introduced and explained. I use four accounting and finance text books in my discussion below. These are i) Palepu et al. (2015) , ii) White, Sondhi, and Fried (2003) iii) Gibson (2007), and iv) Petersen and Plenborg (2012) and I discuss them in turn below.

7.1 Palepu et al. (2015)

> Of course, if one were interested in predicting distress, there would be no need to restrict attention to one variable at a time. A number of multi-factor models have been designed to predict financial distress. One such model, the Altman Z-score model, weights five variables to compute a bankruptcy score.[9] For public companies the model is as follows:[10]
>
> $$Z = 1.2(X_1) + 1.4(X_2) + 3.3(X_3) + 0.6(X_4) + 1.0(X_5)$$
>
> where
>
> X_1 = net working capital/total assets
> X_2 = retained earnings/total assets
> X_3 = EBIT/total assets
> X_4 = market value of equity/book value of total liabilities
> X_5 = sales/total assets
>
> The model predicts bankruptcy when $Z < 1.81$. The range between 1.81 and 2.67 is labelled the 'grey area'.

> Such models have some ability to predict failing and surviving organisations. Altman (1968) reports that, when the model was applied to a holdout sample containing 33 failed and 33 non-failed organisations (the same proportion used to estimate the model), it correctly predicted the outcome in 63 of 66 cases. However, the performance of the model would degrade substantially if applied to a holdout sample where the proportion of failed and non-failed organisations was not forced to be the same as that used to estimate the model.
>
> Several attempts have been made to produce financial distress prediction models using non-US data.[11] Ferner and Hamilton (1987) constructed a New Zealand version of the Altman

Figure 7-1 **From Business analysis & valuation, p. 324, by Palepu et al. (2015)**. Cengage Learning Australia (Permission has been requested to reproduce).

The discussion begins with introducing the Z-score model as a multi-factor model that has been designed to predict company failure. It mentions that the model can be used for "public companies". This is misleading as the original Altman 1968 model is limited to public manufacturing companies; therefore, it is unclear to what degree the model is applicable outside the manufacturing industry. In the 2nd paragraph of Figure 7-1, the authors make an important point in that the sample selection will strongly influence the model usefulness. On the other hand, the authors could emphasise the outdated data: Altman's Z-score model was formulated about 70 years ago and the sample was chosen for a 20 year period from 1945 to 1965. There have been regulation and economic changes since, which also will question the usefulness of the model today.

model; that is, using the same variables but recalibrating the coefficients. The Ferner-Hamilton model is:

$$Z = 0.86 - 3.55(X_1) + 0.5(X_2) + 8.77(X_3) + 0.95(X_4) - 0.99(X_5)$$

Table 10.6 presents calculations of the Ferner-Hamilton model for two New Zealand companies: Feltex and Cavalier.

TABLE 10.6 Ferner-Hamilton model applied to Feltex and Cavalier

Factors	Co-efficients	Feltex Ratios	Feltex Score	Cavalier Ratios	Cavalier Score
Intercept	0.86	0.23	0.86	0.38	0.86
Net working capital/assets	–3.55	0.00	–0.81	0.44	–1.37
Retained earnings/total assets	0.50	0.09	0.00	0.17	0.22
EBIT/total assets	8.77	0.87	0.78	3.47	1.46
Market value of equity/book value of total liabilities	0.95	1.11	0.83	1.44	3.30
Sales/total assets	–0.99	0.23	–1.10	0.38	–1.42
Z-score			–0.29		2.19

Figure 7-2 **From Business analysis & valuation, p. 325**, by Palepu et al. (2015), Cengage Learning Australia (Permission has been requested to reproduce).

In the next section, the authors expose the reader to an alternative Z-score model for New Zealand manufacturing firms reported by Ferner and Hamilton (1987) (cf. Figure 7-2). This is a useful addition that demonstrates there can be a difference in the models *and* the coefficients *and* the intercept. In my thesis above, I have contrasted the addition and the change of the equation parameters in different models to show that when the Z-score is re-estimated it changes.

However, there is an issues in the illustrated example shown in Table 10.6 (cf. Figure 7-2). The authors have made several errors in the table and misrepresented Ferner and Hamilton (1987). The correct table is given in Table 7-1.

Table 7-1 **replication of the Illustrated example in Table 10-6 in Figure 7-2.** From Business analysis & valuation, p. 325, by Palepu et al. (2015), Cengage Learning Australia

Factors	Co-efficients	Feltex **Ratios**	Feltex Score 1	Cavalier **Ratios**	Cavalier Score 2
intercept	0.86	***1.00***	0.86	***1.00***	0.86
X1	-3.55	***0.23***	-0.81	***0.39***	-1.37
X2	0.5	***0.00***	0	***0.44***	0.22
X3	8.77	***0.09***	0.78	***0.17***	1.46
X4	0.95	***0.87***	0.83	***3.47***	3.3
X5	-0.99	***1.11***	-1.1	***1.43***	-1.42
Z-score			**0.56**		**3.05**

I noticed that a ratio was shown for the intercept which is incorrect. Also the individual variable contributions are incorrect and they do not add up to the Z-scores at the bottom for both firms, Feltex and Calvier. The correct Z-scores are given in Table 7-1. For Feltex the figure showed a Z-score of -0.29 whereas in Table 7-1 it calculates to be 0.56 and for Cavalier the Z-scores is 2.19 in Table 7-1 but it should be 3.05. The change of the Z-scores was due to the intercept. Palepu et al. have not included the intercept: when one deducts the intercept of 0.86 from 0.56 and 3.05 it obtains -0.29 and 2.19.

The text does neither explicitly state that the intercept was not included in the analysis nor is a reason provided why it has not been considered. Given the classification dependence on the cut-off values, students would benefit in knowing the influence of adding the intercept. It certainly will influence the classification performance. This very point is made in the next paragraph (cf. Figure 7-3) in which Feltex is used as an example of a failed firm, as allegedly its Z-score of -0.29 is smaller than the cut-off value of 0.04. In fact Feltex would be a misclassification example in that the corrected Z-score is 0.56.

The above Z-scores were estimated from the 2005 financial statements. According to Ferner and Hamilton the Z score that minimises total classification errors is −0.04.

Feltex and Cavalier are both carpet manufacturers. Feltex was floated in June 2004 at $1.70 a share. It subsequently revised its earnings forecasts downwards several times and eventually breached its banking covenants. Receivers were appointed in September 2006. The company went into liquidation after the assets and operations were sold. It is, therefore, not surprising that the Z-score of Feltex is below the cut-off point, while that of Cavalier indicates a low likelihood of failure. Note, also, that it is the market equity to book debt ratio (X_4) that is the major contributor to the differences in Z-score between Feltex and Cavalier. This might be expected because the market price impounds future expectations, including the probability of financial distress.

Financial ratios can differ over time, between different industries and across different accounting methods. Simple distress prediction models like the Altman model tend to be sample-specific and cannot serve as a replacement for in-depth analysis of the kind discussed throughout this book. However, they do provide a useful reminder of the power of financial statement data to summarise important dimensions of the organisation's performance. In addition, they can be useful for screening large numbers of organisations prior to more in-depth analysis of corporate strategy, management expertise, market position and financial ratio performance. In screening data, the *ranking* of 'Z-scores' is likely to be more important than the specific cut-off values.[12]

Figure 7-3 From Business analysis & valuation, p. 325, by Palepu et al. (2015), Cengage Learning Australia Figure (Permission has been requested to reproduce).

7.2 White et al. (2003)

Altman's Z-Score. The best known of the bankruptcy prediction studies that have withstood the test of time is Altman's (1968) Z-score model. The Z-score is the value resulting from the following discriminant analysis equation:

$$
\begin{aligned}
\mathbf{Z} = {} & \mathbf{1.2 \times Working\ capital / Total\ assets} \\
& \mathbf{+ 1.4 \times Retained\ earnings / Total\ assets} \\
& \mathbf{+ 3.3 \times EBIT / Total\ assets} \\
& \mathbf{+ 0.6 \times Market\ value\ of\ equity / Book\ value\ of\ debt} \\
& \mathbf{+ 1.0 \times Sales / Total\ assets}
\end{aligned}
$$

A Z-score below (above) the critical value of 2.675 signals bankruptcy (solvency). Analysis of the misclassifications resulting from use of this critical value resulted in a more intuitively appealing dichotomy:

> It is concluded that all firms having a Z score of greater than 2.99 clearly fall into the "non-bankrupt" sector, while those firms having a Z below 1.81 are all bankrupt. The area between 1.81 and 2.99 will be defined as the "zone of ignorance" or "gray area" because of the susceptibility to error classification.[13]

The original Z-model was designed for manufacturing firms. Additionally, the model was directly applicable only to publicly traded companies because one of its inputs was the market value of equity. To remedy these shortcomings, Altman developed two variations of the Z-model, Z′ and Z″. The Z′-model was developed for nonpublic companies and used the book value of equity in place of the market value of equity. The Z″-model (which omitted the sales turnover ratio) was designed to be applicable to nonmanufacturing (public or pri-

Figure 7-4 **From the analysis and use of financial statements, p. 652**, by White et al., 2003 (Permission has been requested to reproduce).

In this text the authors introduce the Z-score model by putting the equation as shown in Figure 7-4). The text highlights that the model was designed using manufacturing firms and that it is only applicable to publicly traded companies. Furthermore, the text mentions two revised models for private sector and service sector companies. This is a useful contribution: it shows the difference between the models and the cut-off values.

The Z'-score, which is used in this research too, used the book value of equity instead of the market value of equity. The Z"-model has sales turnover removed as compared to the Z'-model. The comparison of these models shows that depending on the sample criteria the model would change.

BOX 18-2

Z′ and Z″: Variations of Altman's Z-Score*

For Private Firms $Z' =$	For Service Sector $Z'' =$
$0.717 \times \frac{\text{Working capital}}{\text{Total assets}}$	$6.56 \times \frac{\text{Working capital}}{\text{Total assets}}$
$+\ 0.847 \times \frac{\text{Retained earnings}}{\text{Total assets}}$	$+\ 3.26 \times \frac{\text{Retained earnings}}{\text{Total assets}}$
$+\ 3.107 \times \frac{\text{EBIT}}{\text{Total assets}}$	$+\ 6.72 \times \frac{\text{EBIT}}{\text{Total assets}}$
$+\ 0.420 \times \frac{\text{Book value of equity}}{\text{Book value of debt}}$	$+\ 1.05 \times \frac{\text{Book value of equity}}{\text{Book value of debt}}$
$+\ 0.998 \times \frac{\text{Sales}}{\text{Total assets}}$	

Appropriate cutoff points for bankruptcy/nonbankruptcy and the gray areas are:

Z′ score	Indication	Z″ score	Indication
<1.23	Bankruptcy	<1.10	Bankruptcy
1.23–2.90	Gray area	1.10–2.60	Gray area
>2.90	Nonbankruptcy	>2.60	Nonbankruptcy

*See Chapter 8 of Edward I. Altman, *Corporate Financial Distress and Bankruptcy* (New York: Wiley, 1993).

Figure 7-5 **From the analysis and use of financial statements, p. 653**, by White et al., 2003, Willey USA (Permission has been requested to reproduce).

One factor that should be noted is the age factor of these models which has been omitted.

7.3 Gibson (2007)

The textbook by Gibson (2007) (cf. Figure 7-6) discusses Altman's Z-score model and the variables therein. The text discusses the complications in the original Z-model (Altman, 1968) of mixing absolute terms with the Sales/Total Asset ratio that was expressed in percentage form. This discussion unnecessarily confuses. On the positive side, the age factor of the model was highlighted, furthermore they mention the differences made for the accounting standards in 1970 and 2005. The text also provides an example: manufacturer Nike and their financial data are used in calculating its Z-score for 2005.

Multivariate Model

Edward I. Altman developed a multivariate model to predict bankruptcy.[10] His model uses five financial ratios weighted in order to maximize the predictive power of the model. The model produces an overall discriminant score, called a Z score. The Altman model is as follows:

$$Z = .012\ X_1 + .014\ X_2 + .033\ X_3 + .006\ X_4 + .010\ X_5$$

$$X_1 = \text{Working Capital/Total Assets}$$

This computation is a measure of the net liquid assets of the firm relative to the total capitalization.

$$X_2 = \text{Retained Earnings (balance sheet)/Total Assets}$$

This variable measures cumulative profitability over time.

$$X_3 = \text{Earnings Before Interest and Taxes/Total Assets}$$

This variable measures the productivity of the firm's assets, abstracting any tax or leverage factors.

$$X_4 = \text{Market Value of Equity/Book Value of Total Debt}$$

This variable measures how much the firm's assets can decline in value before the liabilities exceed the assets and the firm becomes insolvent. Equity is measured by the combined market value of all shares of stock, preferred and common, while debt includes both current and long-term debts.

$$X_5 = \text{Sales/Total Assets}$$

This variable measures the sales-generating ability of the firm's assets.

When computing the Z score, the ratios are expressed in absolute percentage terms. Thus, X_1 (working capital/total assets) of 25% is noted as 25.

The Altman model was developed using manufacturing companies whose asset size was between $1 million and $25 million. The original sample by Altman and the test samples used the period 1946–1965. The model's accuracy in predicting bankruptcies in more recent years (1970–1973) was reported in a 1974 article.[11] Not all of the companies included in the test were manufacturing companies, although the model was initially developed by using only manufacturing companies.

With the Altman model, the lower the Z score, the more likely that the firm will go bankrupt. By computing the Z score for a firm over several years, it can be determined if the firm is moving toward a more likely or less likely position in regard to bankruptcy. In a later study that covered the period 1970–1973, a Z score of 2.675 was established as a practical cutoff point. Firms that scored below 2.675 are assumed to have characteristics similar to those of past failures.[12] Current GAAP recognizes more liabilities than the GAAP used at the time of this study. Thus, we would expect firms to score somewhat less than in the time period 1970–1973. The Altman model is substantially less significant if there is no firm market value for the stock (preferred and common), because variable X_4 in the model requires that the market value of the stock be determined.

The Z score for Nike for 2005 follows:

$$\begin{aligned} Z = {} & .012 \text{ (working capital/total assets)} \\ & + .014 \text{ (retained earnings [balance sheet]/total assets)} \\ & + .033 \text{ (earnings before interest and taxes/total assets)} \\ & + .006 \text{ (market value of equity/book value of total debt)} \\ & + .010 \text{ (sales/total assets)} \end{aligned}$$

$$\begin{aligned} Z = {} & .012\ ([\$6{,}351{,}100{,}000 - \$1{,}999{,}200{,}000])/\$8{,}793{,}600{,}000 \\ & + .014\ (\$4{,}396{,}500{,}000/\$8{,}793{,}600{,}000) \\ & + .033\ (\$1{,}859{,}800{,}000 + \$4{,}800{,}000)/\$8{,}793{,}600{,}000 \\ & + .006\ ([\$261{,}100{,}000 \times \$82.20])/\$3{,}149{,}400{,}000 \\ & + .010\ (\$13{,}739{,}700{,}000/\$8{,}793{,}600{,}000) \end{aligned}$$

$$\begin{aligned} Z = {} & .012\ (49.49) \\ & + .014\ (50.00) \\ & + .033\ (21.20) \\ & + .006\ (681.48) \\ & + .010\ (156.25) \end{aligned}$$

$$Z = .59 + .70 + .70 + 4.09 + 1.56$$

$$Z = 7.64$$

The Z score for Nike for 2005 was 7.64. Considering that higher scores are better and that companies with scores below 2.675 are assumed to have characteristics similar to those of past failures, Nike is a very healthy company.

There are many academic studies on the use of ratios to forecast financial failure. These studies help substantiate that firms with weak ratios are more likely to go bankrupt than firms with strong ratios. Since no conclusive model has yet been developed, the best approach is probably an integrated one. As a supplemental measure, it may also be helpful to compute some of the ratios that appear useful in forecasting financial failure.

Figure 7-6 **From financial reporting analysis, p.491-492,** by Gibson (2007), Cengage Learning USA (Permission has been requested to reproduce).

7.4 Petersen and Plenborg (2012)

Multiple discriminant analysis

Multiple discriminant analysis is a statistical technique used to classify an observation into one of several *a priori* groupings – in this case bankruptcy versus non-bankruptcy groups. A multiple discriminant analysis attempts to derive a linear combination of financial ratios which best distinguish between the bankruptcy and non-bankruptcy firms. Based on a sample of firms that went bankrupt in the past and a random sample of firms that did not, the analysis determines a set of discriminant coefficients. These coefficients are then multiplied on selected financial ratios to obtain a so-called Z-score.

Altman (1968) find that from the original list of 22 financial ratios, five are selected as doing the best job in combination of predicting corporate bankruptcy. Altman estimates the following coefficients on each of the five financial ratios:

$$\text{Z-score} = 1.2\left(\frac{\text{Working capital}}{\text{Total assets}}\right) + 1.4\left(\frac{\text{Retained earnings}}{\text{Total assets}}\right)$$

$$+ 3.3\left(\frac{\text{EBIT}}{\text{Total assets}}\right) + 0.6\left(\frac{\text{Market value of equity}}{\text{Book value of liabilities}}\right) + 1.0\left(\frac{\text{Sales}}{\text{Total assets}}\right)$$

Altman (1968) finds that firms obtaining a Z-score below 1.81 have a high probability of going bankrupt. On the other hand, firms with a Z-score above 2.99 have a low probability of going bankrupt. Firms with a Z-score in between 1.81 and 2.99 are in a grey area, and therefore need to be analysed further.

The Z-score classifies 95% of the observations correctly one year prior to the bankruptcy. There is a 6% chance of a type 1 error and a 3% chance of a type 2 error one year prior the bankruptcy, which supports that the model is doing fairly well. The model has subsequently been further improved (Altman 1983).

Figure 7-7 **From financial statement analysis, p.293-294, by Petersen and Plenborg (2012).** Pearson UK (Permission has been requested to reproduce).

The textbook by Petersen and Plenborg (2012) (cf. Figure 7-7) shows the Z-score model and its cutt-off ranges, and it also metions misclasification rates for Type I and Type II errors for the original 1968 model. Further, the text describes that the model is doing “fairly well”. The bankruptcy literature has, by 2012, shown that the Z-score can perform quite badly (e.g. Bellovary et al., 2007). Overall this textbook does not explain the bankruptcy model in great detail which increases the risks that a readers may use the equation inappropriately.

7.5 Proposed version of the Z-score presentation

The short discussion provided above on the four textbook examples shows that the texts would benefit from adding the important assumptions and limitations of the Z-score model. Because the bankruptcy literature does not have an underlying theory about variable and model selection, it is important to clearly articulate the model's assumptions and limitations. If the model are accompanied by exhibits in the textbook, these should explain interdependence between the measures and the cut-off values.

Below I draft my suggestion about how to present the Z-score model in an accounting textbook.

Bankruptcy prediction using multivariate discriminant analysis

Altman's Z-score (Altman, 1968) is a popular bankruptcy prediction model. The Z-score is a special case from the following generic equation

$$Z = \alpha + \beta_1\, X1 + \beta_2\, X2 + \beta_3\, X3 + \beta_4\, X4 + \cdots + \beta_n\, Xn. \quad (1)$$

The size of the weightings α and the β_is (*i =1, 2... n)* depend on a sample of firms which the user is free to choose. Similarly, it is up to the user to choose the type and quantity of financial ratios X*i* (*i*=1,2,.n). Altman's (1968) choice was as follows: 66 US manufacturing firms of which 33 had failed and 33 had not failed between 1945 and 1965; 5 financial ratios. The combination of both choices resulted in

$$Z = 1.0\ \frac{Sales}{Total\ Assets} + 1.4\ \frac{Retained\ Earnings}{Total\ Assets} + 3.3\ \frac{EBIT}{Total\ Assets} + 0.6\ \frac{Market\ Value}{Total\ Liabilities} + 1.2\ \frac{Working\ Capital}{Total\ Assets} \quad (2)$$

and an equation-specific classification rule: for a firm, if the weighted Z-index is Z > 2.99, it is classified as healthy and if Z < 1.81 it is classified as failed, and if it is in between the two values, 1.81< Z < 2.99, no clear healthy-failed group assignment can be made…

In the above draft I have introduced the generic Z-score model to demonstrate that the equation parameters and cut-off values in the original model depend on the estimation sample. I have given the reader information about the sample, variables and the data chosen to estimates Equation (2) in which I alert them of time-sensitivity.

Following this first few paragraphs, the focus may then be selected. The additional information that should be included would describe some of the restrictions in using the original variables, because they are not universally representative of all failure modes. If the reader is interested in estimating the model, then information on the model usefulness, such as model predictive ability, sample and the model assumptions would be useful. The predictive ability of Altman's (1968) Z-score depends on his classification rule, where he obtained two cutt-off values into which one can classify firms (failed and non-failed). Classifications into two of the groups can be controlled for, and two types of errors can be made, Type I and Type II. It then would be useful to elaborate the costliness of each error because they usually are different in size.

Although the detail of the analysis on the Z-score would depend on the amount of the words and the space the author of a textbook wishes to use, the following minimal information should be included:

- Predictive performance and the assumptions around the classification rule. Have the readers been made aware of the Type I and Type II error classifications and the costs.
- Re-estimations and the variable usefulness. Have the reader being made aware of limitations of using the original Z-score variables. They do not capture all failure modes.

Chapter 8 Research summary and conclusion

Motivated by the recent Carillion PLC collapse, I have investigated the applicability of Altman's (1968) Z-score model and extensions thereof to the construction industry. I have investigated whether adding AFT and SMT, two qualitative variables, would improve the forecasting performance of the multivariate linear discriminant model. The two variables are initially measured on a nominal scale. Based on the literature, I have operationalised these onto a semi-interval (ordinal with possibly unequal distances between the groups) scale in order for them to be included in the linear discriminant regressions. I've searched extensively for a best parametrization of the distances to yield 6- and 7-variables Z-score models. All models have been tested for error-term behaviour (one important statistical model assumption) and predictive performance (model usefulness).

The misclassification rates were reduced by the addition of the AFT and SMT variable in comparison to an extended (intercept) and re-estimated Altman Z-score. The predictive accuracy of the 6- and 7-variable models (Z'_{SMT}, Z'_{AFT} and $Z'_{AFT,SMT}$) showed a robust and steady performance for all reasonable scale distances. However, there was a negative correlation between the error term Normality and the predictive performance – one would expect otherwise because error non-Normality suggests an omitted variables problem which would reduce the number of different failure modes the model is able to detect, hence increase the misclassification rate. The findings show that the addition of the two qualitative variables increased the prediction accuracy in both the estimation (matched-pair with N=40) and test (matched-pair N=30) samples. The best model by predictive accuracy is the 7-variables model: it predicts 98.6% of failed and non-failed companies in my full sample of 70 firms (i.e. 69-out-of-70) one year ahead.

I then have applied the knowledge gained to the case of Carillion PLC and the critically appreciated the presentation of the Z-score model in accounting textbooks.

All of my final 6- and 7-variables models predict the failure of Carillion at least 2 years ahead. Altman's re-estimated and intercept-including 5-variables model did not – it showed that Carillion operates poorly over a decade but with an upwards trend in the Z-scores prior to the company's failure. Secondly, during my research, I gained an appreciation of the sensitivity of coefficient estimates and cut-off values to the choice of the estimation sample and functional form of the model. These issues then translate into a model's prediction performance. These issues I believe would benefit readers of accounting textbooks when they are about to make investment decisions based on the original 1968 Altman Z-score.

8.1 Research Limitations and future directions

There are several limitations in this research which in turn provide for future research avenues. Firstly, I used a discrete set of 136 cases to develop three measurement distance scales for my two qualitative ordinal variables and their joint use. However, in an extended Monte Carlo simulation I would run different estimation-test sample compositions using the 70 companies against an encompassing domain of input distance permutations (representing the three discrete and ordered groups G1, G2 and G3).

Secondly, I chose a matched sample for failed and non-failed companies, which may not represent entry and exit frequecies observed in the industry. Thus, if there are more construction industry entrants, the matched sample over-represents the failed companies which will bias the parameter estimates.

Chapter 9 References

Adams, T., Krishnan, J., & Krishnan, J. (2017). *Client influence and auditor independence revisited: evidence from auditor resignations*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3024877

Agrawal, A., Jaffe, J. F., & Karpoff, J. M. (1999). Management turnover and governance changes following the revelation of fraud. *The Journal of Law and Economics, 42*(S1), 309-342.

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(3), 370-374.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589-609.

Altman, E. I. (1983). *Corporate financial distress: a complete guide to predicting, avoiding, and dealing with bankruptcy*. New York: Wiley.

Altman, E. I. (2000). Predicting financial distress of companies: revisiting the Z-score and ZETA models. *Stern School of Business, New York University*, 9-12. https://doi.org/10.4337/9780857936097.00027

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of altman's Z-Score model. *Journal of International Financial Management & Accounting, 28*(2), 131-171. https://doi.org/10.1111/jifm.12053

Argenti, J. (1976). Corporate planning and corporate collapse. *Long Range Planning, 9*(6), 12-17. https://doi.org/10.1016/0024-6301(76)90006-6

Arrunada, B., & Paz-Ares, C. (1997). Mandatory rotation of company auditors: A critical examination. *International Review of Law and Economics, 17*(1), 31-61.

Assaad, K. W. (2010). *Financial statements as sole predictors of financial distress: The need for new financial ratios* (Masters

University of Neuchâtel, Switzerland). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1797852

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting Review, 38*(1), 63-93.

Beattie, V., & Fearnley, S. (1995). The importance of audit firm characteristics and the drivers of auditor change in UK listed companies. *Accounting and Business Research, 25*(100), 227-239.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 71-111.

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial education*, 1-42.

Blay, A. D., Geiger, M. A., & North, D. S. (2011). The auditor's going-concern opinion as a communication of risk. *Auditing: A Journal of Practice & Theory, 30*(2), 77-102.

Carey, P., & Simnett, R. (2006). Audit partner tenure and audit quality. *The Accounting Review, 81*(3), 653-676.

Chen, K. C., & Church, B. K. (1996). Going concern opinions and the market's reaction to bankruptcy filings. *Accounting Review*, 117-128.

Comeptition watchdog recommends accountancy market overhaul. (2019). *BBC*. Retrieved from https://www.bbc.com/news/business-47969528

Darrat, A. F., Gray, S., Park, J. C., & Wu, Y. (2016). Corporate governance and bankruptcy risk. *Journal of Accounting, Auditing & Finance, 31*(2), 163-202.

Dart, E. (2011). UK investors' perceptions of auditor independence. *The British Accounting Review, 43*(3), 173-185.

Davies, R. (2018). Carillion: Accountants accused of 'feasting' on company. *The Gardian*. Retrieved from https://www.theguardian.com/business/2018/feb/13/carillion-accountants-accused-of-feasting-on-company

Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research*, 167-179.

DeFond, M. L., & Park, C. W. (2001). The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review, 76*(3), 375-404.

du Jardin, P. (2009). Bankruptcy prediction models: How to choose the most relevant variables? *Bankers, Markets & Investors*(98,), 39-46.

Edmister, R. O. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis, 7*(2), 1477-1493.

Ferner, D., & Hamilton, R. (1987). A note on the predictability of financial distress in New Zealand listed companies. *Accounting & Finance, 27*(1), 55-63.

Firth, M. (1981). Auditor-client relationships and their impact on bankers' perceived lending decisions. *Accounting and Business Research, 11*(43), 179-188.

Furtado, E. P., & Karan, V. (1990). Causes, consequences, and shareholder wealth effects of management turnover: A review of the empirical evidence. *Financial Management*, 60-75.

Ghosh, A., & Tang, C. Y. (2015). Auditor resignation and risk factors. *Accounting Horizons, 29*(3), 529-549.

Gibson, C. H. (2007). *Financial reporting & analysis: using financial accounting information* (10th Internat. student ed.). Mason, Ohio: South-Western.

Gilson, S. C. (1989). Management turnover and financial distress. *Journal of financial Economics, 25*(2), 241-262.

Hair, J. F. (1998). *Multivariate data analysis*. Upper Saddle River, N.J: Prentice Hall.

Hinchliffe, R. (2010, 29 January). HBF Wales weekly news summary Retrieved from https://www.hbf.co.uk/articles/news/weekly-news-summary-all/wales-editions/hbf-wales-weekly-news-summary-friday-29-january-2010/

Hoyle, J. (1978). Mandatory auditor rotation: The arguments and an alternative. *Journal of Accountancy (pre-1986), 145*(000005), 69.

Hsu, H.-H., & Wu, C. Y.-H. (2014). Board composition, grey directors and corporate failure in the UK. *The British Accounting Review, 46*(3), 215-227.

Hudaib, M., & Cooke, T. (2005). The impact of managing director changes and financial distress on audit qualification and auditor switching. *Journal of Business Finance & Accounting, 32*(9-10), 1703-1739.

Hussey, R. (1999). The familiarity threat and auditor independence. *Corporate Governance: An International Review, 7*(2), 190-197.

ICAEW. (2019). APB amends rotation requirements for audit partners on listed entity audits. Retrieved from https://www.icaew.com/technical/ethics/auditor-independence/apb-amends-rotation-requirements-for-audit-partners-on-listed-entity-audits

IESBA. (2012). Strengthening safegurds against familiarity threats, 12. Retrieved from International Ethics Standards Board for Accountants website: https://www.ethicsboard.org/system/files/meetings/files/IESBA%20-%20Familiarity%20Threats%20-CAG%20agenda%20paper%20E-1.pdf

Kaplan, S. N. (1993). *Top executives, turnover and firm performance in Germany*.National Bureau of Economic Research.

Knapp, M. C. (1991). Factors that audit committee members use as surrogates for audit quality. *Auditing-A Journal of Practice & Theory, 10*(1), 35-52.

Laitinen, E. K. (1992). Prediction of failure of a newly founded firm. *Journal of Business Venturing, 7*(4), 323-340.

Lindahl, F. (1992). *How much do we know about auditor changes? The effect of observed and unobserved heterogeneity'*. Paper presented at the AAA Annual Conference.

Marriage, M. (2018). Probe urged into break-up of Big four accountants. Retrieved from https://www.ft.com/content/911e8184-283d-11e8-b27e-cc62a39d57a0

Mason, R., & Harris, F. (1979). Predicting company failure in the construction industry. *Proceedings of the Institution of Civil Engineers, 66*(2), 301-307.

McLaren, N. (1958). Rotation of auditors. *Journal of Accountancy, 106*(1), 41-44.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 109-131.

Ono, M. (2018). *Comparing the predictive performance of the logit and discriminant models in bankruptcy prediction*. Department of Acounting and Information systems, University of Canterbury.

Palepu, K. G., Healy, P. M., Bernard, V. L., Wright, S., Bradbury, M., & Lee, P. (2015). *Business analysis & valuation: using financial statements* (Second Asia Pacific ed.). South Melbourne, Victoria: Cengage Learning Australia.

Pervan, I., & Kuvek, T. (2013). The relative importance of financial ratios and nonfinancial variables in predicting of insolvency. *Croatian Operational Research Review, 4*(1), 187-197.

Petersen, C. V., & Plenborg, T. (2012). *Financial statement analysis: valuation, credit analysis, executive compensation*. Harlow [u.a.]: Financial Times, Prentice Hall/Pearson.

Schwartz, K. B., & Menon, K. (1985). Executive succession in failing firms. *Academy of Management Journal, 28*(3), 680-686.

Sikka, P. (1992). Audit policy making in the UK: The case of 'the auditor's considerations in respect of going concern'. *European Accounting Review, 1*(2), 349-392.

Taffler, R. J. (1982). Forecasting company failure in the UK using discriminant analysis and financial ratio data. *Journal of the Royal Statistical Society: Series A (General), 145*(3), 342-358.

Wearmouth, R. (2018). Carillion directors ripped apart by MPs in scathing report. *Huffpost*. Retrieved from https://www.huffingtonpost.co.uk/entry/too-busy-stuffing-their-mouths-with-gold-mps-rip-apart-carillion-directors-in-scathing-report_uk_5afaf28ae4b0200bcab90272

White, G. I., Sondhi, A. C., & Fried, D. (2003). *The analysis and use of financial statements* (3rd ed.). New York: Wiley.

Wissmann, M., & Toutenburg, H. (2007). *Role of categorical variables in multicollinearity in the linear regression model*. Department of Statistics: Technical Reports: Retrieved from https://epub.ub.uni-muenchen.de/2081/1/report008_statistics.pdf

Appendix

Table A-1 Depicts the chosen 136 cases to run the Monte Carlo Simulation. The simulation is run using each case to measure the AFT/SMT variable and plugging into $Z'_{SMT_{n(1,..,136)}} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_6 SMT$ or $Z'_{AFT_{n(1,..,136)}} = \alpha + \beta_1 STA + \beta_2 RETA + \beta_3 EBITTA + \beta_4 BVETL + \beta_5 WCTA + \beta_6 AFT$ Models. (Chapter 5)

G3 SIMULATION				G2 SIMULATION				G1 SIMULATION			
Case	G3	G2	G1	Case	G3	G2	G1	Case	G3	G2	G1
SMT/AFT1	-0.9	-0.3	0.9	SMT/AFT42	-0.9	-0.3	0.9	SMT/AFT89	-0.9	-0.3	0.9
SMT/AFT2	-0.9	-0.2	0.8	SMT/AFT43	-0.8	-0.3	0.8	SMT/AFT90	-0.8	-0.2	0.9
SMT/AFT3	-0.9	-0.1	0.7	SMT/AFT44	-0.7	-0.3	0.7	SMT/AFT91	-0.7	-0.1	0.9
SMT/AFT4	-0.9	0	0.6	SMT/AFT45	-0.6	-0.3	0.6	SMT/AFT92	-0.6	0	0.9
SMT/AFT5	-0.9	0.1	0.5	SMT/AFT46	-0.5	-0.3	0.5	SMT/AFT93	-0.5	0.1	0.9
SMT/AFT6	-0.9	0.2	0.4	SMT/AFT47	-0.4	-0.3	0.4	SMT/AFT94	-0.4	0.2	0.9
								SMT/AFT95	-0.3	0.3	0.9
SMT/AFT7	-0.8	-0.3	0.9	SMT/AFT48	-0.9	-0.2	0.9	SMT/AFT96	-0.9	-0.3	0.8
SMT/AFT8	-0.8	-0.2	0.8	SMT/AFT49	-0.8	-0.2	0.8	SMT/AFT97	-0.8	-0.2	0.8
SMT/AFT9	-0.8	-0.1	0.7	SMT/AFT50	-0.7	-0.2	0.7	SMT/AFT98	-0.7	-0.1	0.8
SMT/AFT10	-0.8	0	0.6	SMT/AFT51	-0.6	-0.2	0.6	SMT/AFT99	-0.6	0	0.8
SMT/AFT11	-0.8	0.1	0.5	SMT/AFT52	-0.5	-0.2	0.5	SMT/AFT100	-0.5	0.1	0.8
SMT/AFT12	-0.8	0.2	0.4	SMT/AFT53	-0.4	-0.2	0.4	SMT/AFT101	-0.4	0.2	0.8
				SMT/AFT54	-0.3	-0.2	0.3	SMT/AFT102	-0.3	0.3	0.8
SMT/AFT13	-0.7	-0.3	0.9	SMT/AFT55	-0.9	-0.1	0.9	SMT/AFT103	-0.9	-0.3	0.7
SMT/AFT14	-0.7	-0.2	0.8	SMT/AFT56	-0.8	-0.1	0.8	SMT/AFT104	-0.8	-0.2	0.7
SMT/AFT15	-0.7	-0.1	0.7	SMT/AFT57	-0.7	-0.1	0.7	SMT/AFT105	-0.7	-0.1	0.7
SMT/AFT16	-0.7	0	0.6	SMT/AFT58	-0.6	-0.1	0.6	SMT/AFT106	-0.6	0	0.7
SMT/AFT17	-0.7	0.1	0.5	SMT/AFT59	-0.5	-0.1	0.5	SMT/AFT107	-0.5	0.1	0.7
SMT/AFT18	-0.7	0.2	0.4	SMT/AFT60	-0.4	-0.1	0.4	SMT/AFT108	-0.4	0.2	0.7
				SMT/AFT61	-0.3	-0.1	0.3	SMT/AFT109	-0.3	0.3	0.7
SMT/AFT19	-0.6	-0.3	0.9	SMT/AFT62	-0.9	0	0.9	SMT/AFT110	-0.9	-0.3	0.6
SMT/AFT20	-0.6	-0.2	0.8	SMT/AFT63	-0.8	0	0.8	SMT/AFT111	-0.8	-0.2	0.6
SMT/AFT21	-0.6	-0.1	0.7	SMT/AFT64	-0.7	0	0.7	SMT/AFT112	-0.7	-0.1	0.6
SMT/AFT22	-0.6	0	0.6	SMT/AFT65	-0.6	0	0.6	SMT/AFT113	-0.6	0	0.6
SMT/AFT23	-0.6	0.1	0.5	SMT/AFT66	-0.5	0	0.5	SMT/AFT114	-0.5	0.1	0.6
SMT/AFT24	-0.6	0.2	0.4	SMT/AFT67	-0.4	0	0.4	SMT/AFT115	-0.4	0.2	0.6
				SMT/AFT68	-0.3	0	0.3	SMT/AFT116	-0.3	0.3	0.6
SMT/AFT25	-0.5	-0.3	0.9	SMT/AFT69	-0.9	0.1	0.9	SMT/AFT117	-0.9	-0.3	0.5
SMT/AFT26	-0.5	-0.2	0.8	SMT/AFT70	-0.8	0.1	0.8	SMT/AFT118	-0.8	-0.2	0.5
SMT/AFT27	-0.5	-0.1	0.7	SMT/AFT71	-0.7	0.1	0.7	SMT/AFT119	-0.7	-0.1	0.5
SMT/AFT28	-0.5	0	0.6	SMT/AFT72	-0.6	0.1	0.6	SMT/AFT120	-0.6	0	0.5
SMT/AFT29	-0.5	0.1	0.5	SMT/AFT73	-0.5	0.1	0.5	SMT/AFT121	-0.5	0.1	0.5
SMT/AFT30	-0.5	0.2	0.4	SMT/AFT74	-0.4	0.1	0.4	SMT/AFT122	-0.4	0.2	0.5
				SMT/AFT75	-0.3	0.1	0.3	SMT/AFT123	-0.3	0.3	0.5
SMT/AFT31	-0.4	-0.3	0.9	SMT/AFT76	-0.9	0.2	0.9	SMT/AFT124	-0.9	-0.3	0.4
SMT/AFT32	-0.4	-0.2	0.8	SMT/AFT77	-0.8	0.2	0.8	SMT/AFT125	-0.8	-0.2	0.4
SMT/AFT33	-0.4	-0.1	0.7	SMT/AFT78	-0.7	0.2	0.7	SMT/AFT126	-0.7	-0.1	0.4
SMT/AFT34	-0.4	0	0.6	SMT/AFT79	-0.6	0.2	0.6	SMT/AFT127	-0.6	0	0.4
SMT/AFT35	-0.4	0.1	0.5	SMT/AFT80	-0.5	0.2	0.5	SMT/AFT128	-0.5	0.1	0.4
SMT/AFT36	-0.4	0.2	0.4	SMT/AFT81	-0.4	0.2	0.4	SMT/AFT129	-0.4	0.2	0.4
				SMT/AFT82	-0.3	0.2	0.3	SMT/AFT130	-0.3	0.3	0.4
				SMT/AFT83	-0.9	0.3	0.9	SMT/AFT131	-0.9	-0.3	0.3
SMT/AFT37	-0.3	-0.2	0.8	SMT/AFT84	-0.8	0.3	0.8	SMT/AFT132	-0.8	-0.2	0.3
SMT/AFT38	-0.3	-0.1	0.7	SMT/AFT85	-0.7	0.3	0.7	SMT/AFT133	-0.7	-0.1	0.3
SMT/AFT39	-0.3	0	0.6	SMT/AFT86	-0.6	0.3	0.6	SMT/AFT134	-0.6	0	0.3
SMT/AFT40	-0.3	0.1	0.5	SMT/AFT87	-0.5	0.3	0.5	SMT/AFT135	-0.5	0.1	0.3
SMT/AFT41	-0.3	0.2	0.4	SMT/AFT88	-0.4	0.3	0.4	SMT/AFT136	-0.4	0.2	0.3

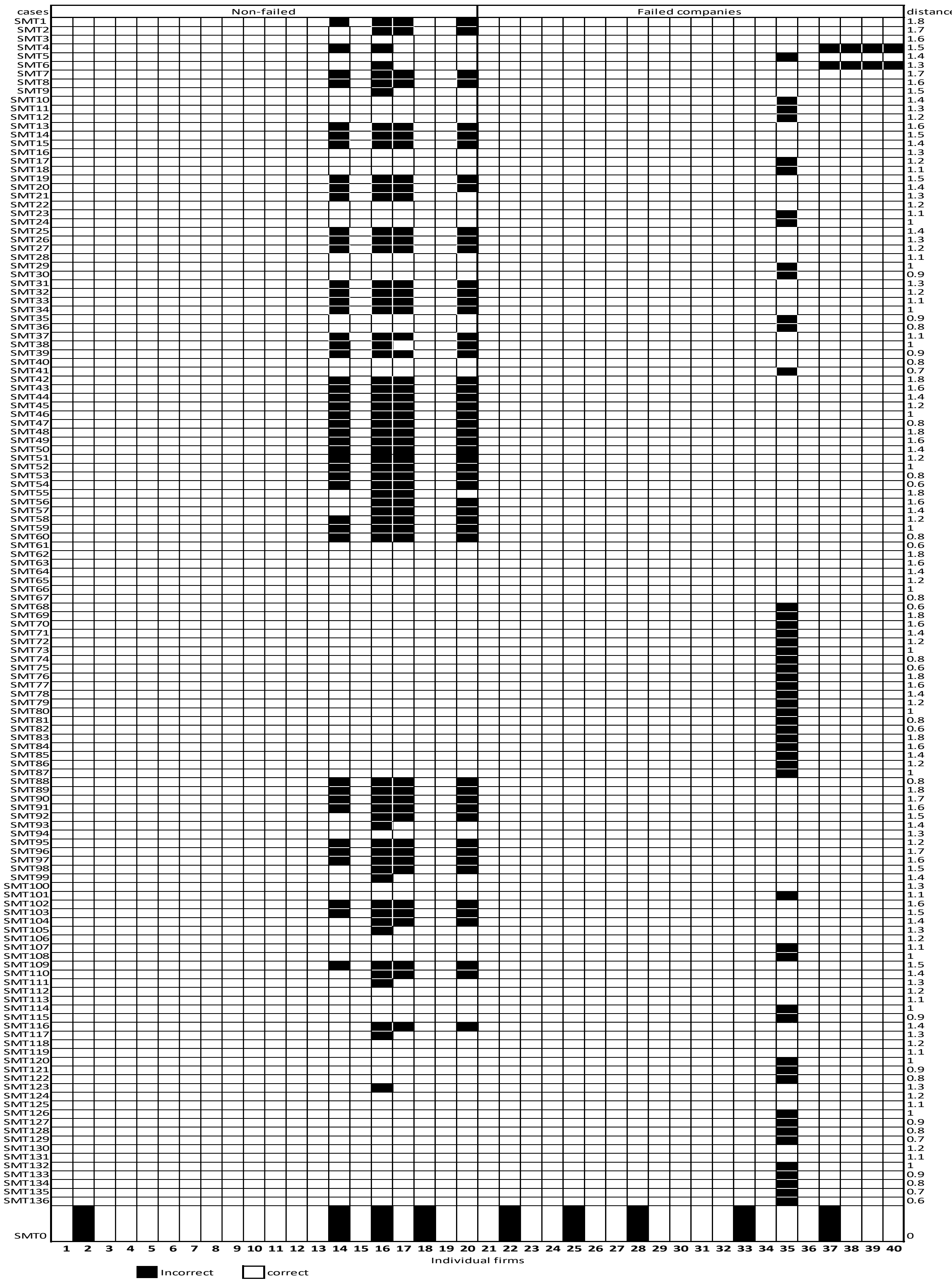


Figure A-1 The individual firm misclassification for all Z'_{SMT} models run for the estimation sample (N=40). SMT0 is the Z' model. (Chapter 5.4)

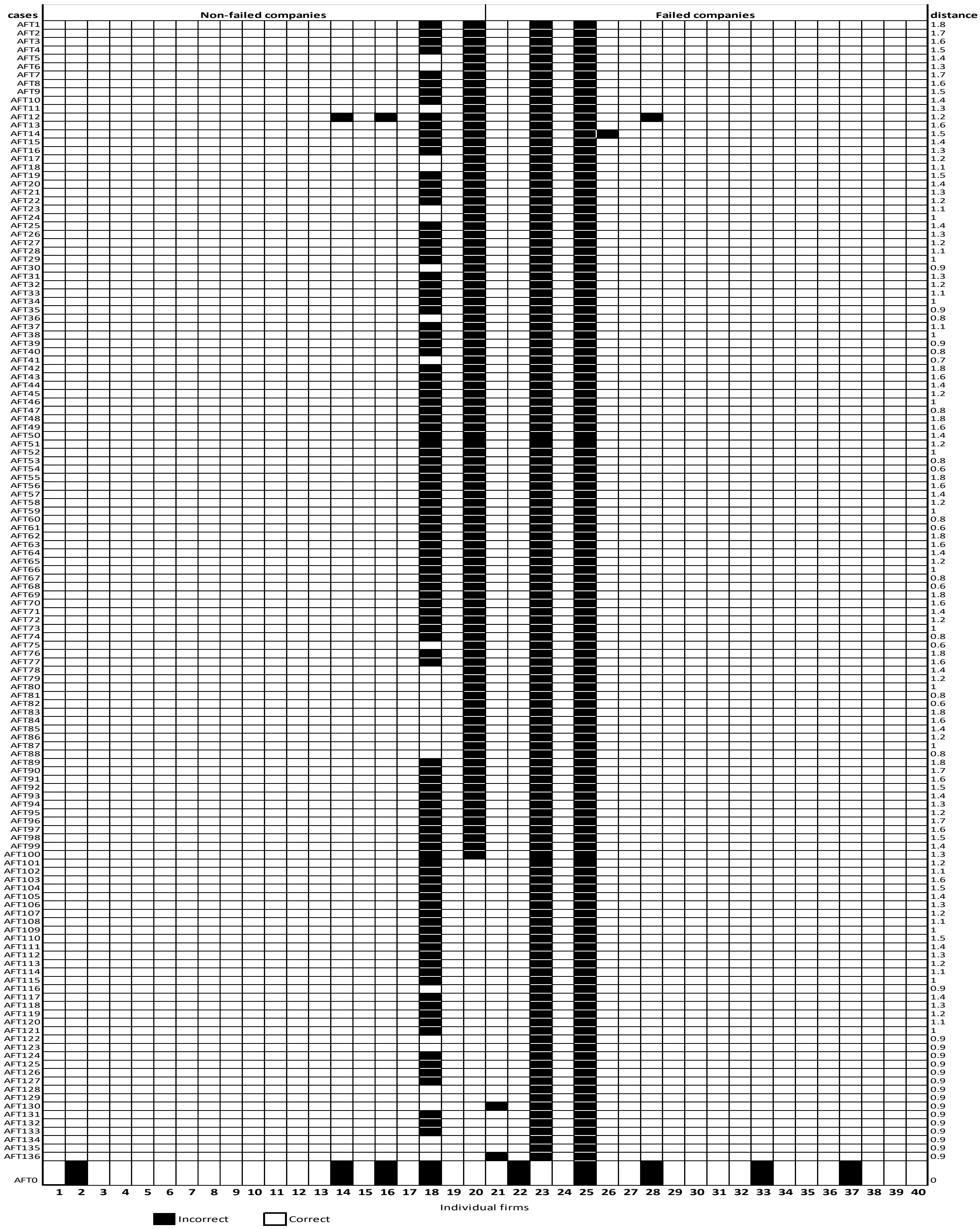


Figure A-2 The individual firm misclassification for all Z'_{AFT} models run for estimation sample. AFT0 is the Z'model. (Chapter 5.4)

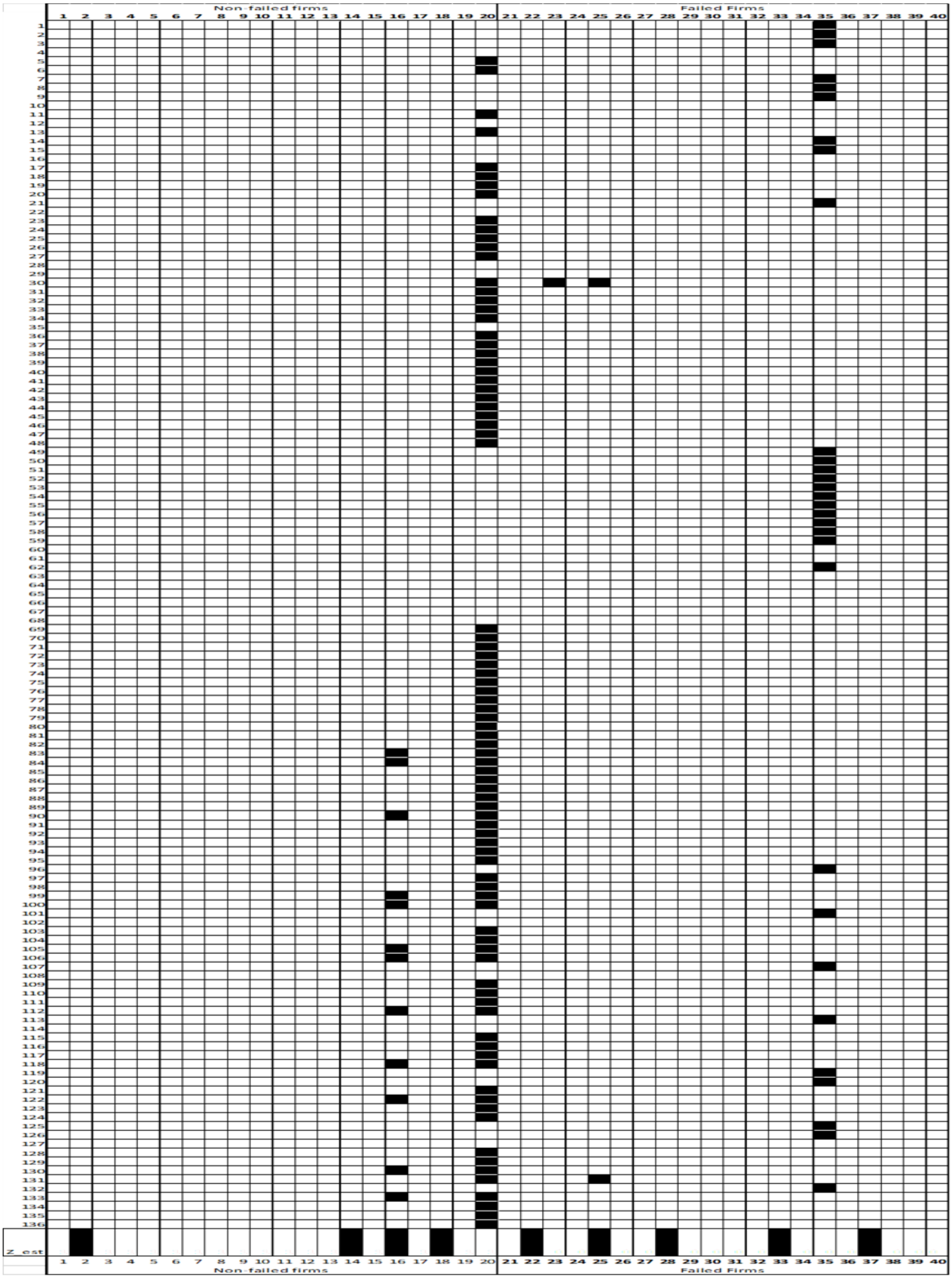


Figure A-3 The individual firm misclassification for the $Z'_{AFT,SMT}$ models run for estimation sample.(Chapter 5.5)